THE UNIVERSITY OF ALBERTA

AN AUTOPSY DATA STORAGE AND RETRIEVAL SYSTEM

by

© MARVIN LIONEL BRAUDE

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT OF COMPUTING SCIENCE

EDMONTON, ALBERTA

FALL, 1971

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled AN AUTOPSY DATA STORAGE AND RETRIEVAL SYSTEM submitted by Marvin Lionel Braude in partial fulfillment of the requirements for the degree of Master of Science.

Date 13 Aug 71

ABSTRACT

This thesis presents a description of The University of Alberta Hospital Autopsy Storage and Retrieval System. The reason for developing this System was to provide the Department of Pathology with a computerized storage and retrieval system for the data obtained from post-mortem examinations. The main purpose of the System is to simplify and shorten the manual procedures presently employed in the compilation of data for research needs; however, it will also greatly reduce the time spent in the compilation of year-end statistics.

Unlike the majority of other systems of this nature, the Autopsy System does not use any special coding of data. The autopsy data is stored in natural language form, and questions submitted for searching are in like form. The most notable expansion over other systems is the ability of the System to search on not only the general information and pathological diagnosis associated with a case, but also the recorded measures, drugs and other administered therapy, as well as the medical history.

ACKNOWLEDGEMENT

I wish to express my sincere appreciation to my supervisor, Dr. D. Fenna for his advice, criticism, and guidance throughout the duration of this research; to Dr. G.O. Bain for his assistance and continued interest in this project; and to Betty Florchyk for an excellent job in typing the final version of this thesis.

I am grateful to the Department of Computing Science for allocating the financial support received in this research.

Special thanks are extended to my wife Pat for her sincere understanding, patience, help, and constant encouragement during the past year.

TABLE OF CONTENTS

	Page
CHAPTER I - INTRODUCTION	1
1.1 Computers in a Hospital Environment	1
1.2 The Problem	1
CHAPTER II - INFORMATION STORAGE AND RETRIEVAL IN PATHOLOGY	3
CHAPTER III - FORMULATION OF A QUESTION	9
3.1 The Question Card	10
3.2 The Term Card	11
3.2.1 Search Logic	12
3.2.2 Search Types	16
3.2.2.1 Name	16
3.2.2.2 Post-Mortem Number	17
3.2.2.2 Age	19
3.2.2.4 Sex	20
3.2.2.5 Race	20
3.2.2.6 Clinical Diagnosis	21
3.2.2.7 Pathological Diagnosis	22
3.2.2.8 Measures	24
3.2.2.9 Drugs and Other Therapy	26
3.2.2.10 Medical History	33
3.3 An Example	34
3.4 Summary	36
CHAPTER IV - QUESTION PROCESSING	38
4.1 Creating the Question File	38
4.2 Error Correction Within a Question	39
4.3 The Searching Phase	42
4.3.1 The SEARPROC File	46
4.4 The Search Output	48
4.4.1 Full Output	48
4.4.2 Abbreviated Output	51
4.5 Batch Mode Versus On-Line Mode	52

	Page
CHAPTER V - THE AUTOPSY DATA BASE	55
5.1 Data Base Structure	57
5.1.1 General Information Division	58
5.1.2 Pathological Diagnosis Division	61
5.1.3 Measures Division	62
5.1.4 Drugs and Other Therapy Division	63
5.1.5 Medical History Division	64
5.1.6 Conclusion of the Data Base	65
5.2 Updating the Autopsy Data Base	66
5.2.1 Updating: A User's View	67
5.2.2 Updating: A System's View	72
5.3 Data Base Error Correction	76
5.3.1 Error Correction: A User's View	76
5.3.2 Error Correction: A System's View	79
CHAPTER VI - THE DATA BASE SEARCHING ALGORITHM	84
6.1 The MAIN Line	84
6.2 Subroutine SEARCH	90
6.3 Subroutine LOCATE	91
6.4 Subroutine COMPAR	91
6.5 Subroutine PAGE	92
CHAPTER VII - SUMMARY	93
BIBLIOGRAPHY	95
APPENDIX	99

LIST OF FIGURES

		Page
Figure 5.1	Updating Procedure	74
5.2	Processing in ERROR	81
5.3	Processing in UPDATE	82
6.1	The TERM Table	85
6.2	The POSITION Table	86
6.3	The QUESTION-PARAMETER Table	88

CHAPTER I

INTRODUCTION

1.1 Computers in a Hospital Environment

In the past decade many new fields of application for the computer have been discovered, one of these being the automation of procedures within a hospital. Hospital applications range from payroll and familiar business operations through the handling of medical records to real-time monitoring of patients and equipment. Hospital computer systems could be used to link outside service bureaus, or as in-house general purpose installations, or dedicated units with highly specialized attachments.

Present visions are to have an integrated network for the purpose of records, communications, and calculations; however, this vision is far off. Presently the vast majority of hospital records, including autopsy data, are stored on paper and searching them is usually very time consuming.

1.2 The Problem

The purpose of the research outlined in this thesis is to provide the Department of Pathology of the University of Alberta Hospital with a storage and retrieval system for the data obtained from post-mortem examinations. This System is a small step towards the realization of future visions as it simplifies and shortens present manual procedures involved in retrieving data for research

purposes. Furthermore, the developed System would be used for the compilation of year-end statistics on autopsies performed.

Chapter II describes previously published research in the field of storage and retrieval of pathological data. Chapters III - V serve as a manual for all aspects of the University of Alberta Hospital Autopsy Storage and Retrieval System. Chapter III outlines the method of formulating a question; Chapter IV outlines the method of submitting a formulated question to the System; Chapter V describes the data base - its structure, method of updating, and method of correcting errors therein. Chapter VI describes the algorithm employed to search the data base and this is followed by a summary in the final chapter.

CHAPTER II

INFORMATION STORAGE AND RETRIEVAL IN PATHOLOGY

The problem of storing and retrieving data from pathological examinations has been one facing pathologists throughout the world for many years. With the advent of digital computers people began realizing the possibility of applying the computer to this pathological problem.

One of the earliest systems for storage and retrieval of data from autopsies was developed by Dr. H.M. Carpenter [5] and termed the 'Peek-a-boo System'. This system makes use of a deck of 1500 cards, each representing a different pathological concept; it does not require a computer. The 1500 cards are divided into three decks of 500 cards and each deck further subdivided into five different colour groups. Each card is assigned a code according to its deck, colour, and sequence number within the deck; each code corresponds to one of the 1500 entries within the vocabulary. Each card is subdivided into a 100 x 100 matrix to allow 10000 different autopsy numbers to be recorded thereon. After pulling the appropriate cards for an autopsy report, a hole, positioned according to the autopsy number, is punched in each card. Retrieval of information from this system involves alignment of the holes on the cards.

The Department of Pathology at Western Reserve University in Cleveland, Ohio [18-22] pioneered research in the field of storage and retrieval of autopsy data by

computer technique. The system developed there permits the storage and retrieval of terminology used to express pathological diagnoses, which is considerably less than the capabilities of the Autopsy System developed for the Department of Pathology at the University of Alberta Hospital. The findings of the pathological diagnosis are keypunched, one word to a card, in a specified format and then transferred to a tape. The next step involves encoding the English words and obtaining a completely coded diagnosis. Codes are assigned by the computer with the aid of a stored dictionary. The coding scheme consists of a series of letters and numbers. The first four letters indicate the category, and this can be followed by two columns of two-digit numbers and two columns of two letters each. The four additional columns allow for divisions within each category and the preservation of word relationships. For example, 'tumor', and all its denotations, is assigned the code PSSA-13; if malignant the code is PSSA-13-11. The first column of letters reveals the composition of the tumor and the second column of letters further subdivides each composition. PSSA-13-11-AA signifies malignant tumors of fibrous connective tissue; PSSA-13-AA-CC specifies only fibroscomas. If a word that is not present in the dictionary is encountered, a diagnostic is generated and facilities exist to allow for its inclusion. Searching the data base consists of encoding the submitted question, which may make use of Boolean logic

to link terms, and performing comparisons against the computer tape. In order to avoid having a built-in decoder, i.e. from the coding scheme to English, or retention of the original uncoded tape, output consists of only the autopsy number, which directs a user to the departmental records.

B.G. Lamson, B. Dimsdale, and H. Jacobs [9-11] developed a partially automated system for the storage and retrieval of surgical pathological data. The system allows the usage of natural English language for the entry of data and posing of questions. All input in natural language is then converted to coded form. Due to the limited size of the vocabulary associated with this particular application, a dictionary of synonym classes linked together by a subordination property was constructed, i.e. a thesaurus. The synonym class permits the use of words with the same meaning as well as the introduction of such items as common misspellings. The searching algorithm permits Boolean AND and NOT operators and is similar to the University of Alberta Hospital Autopsy System in that only one pass through the data base is required to process a batch of questions. The system provides several types of searches - a search based exclusively on the terms requested, a search on synonyms of the terms but not subordinates within the linkage scheme, a search for reports by identifier, and a complex search making use of the compound synonym classes. This system was implemented at the University of California, Los Angeles

and, even though the results were "rather remarkable", it was found to be "rather cumbersome to use". This implementation was on an IBM 7040/7094 directly coupled system, but plans called for the conversion to the System 360 and possibly multiple thesauri.

A.W. Pratt and L.B. Thomas [16] developed a computer-based pathology information processing system which included an autopsy data file, a surgical pathological data file, and a cytopathology data file. Each autopsy data record consists of three introductory lines containing general information about the case, such as hospital registry number and autopsy report number, followed by the final pathological diagnosis. This diagnosis is coded using the Systematized Nomenclature of Pathology (SNOP) and is stored with the coded form and English form side by side. This form of storage seems wasteful since each item of information within the pathological diagnosis is doubly stored. Coding of the data is performed manually and thus a substantial amount of time is wasted prior to keypunching. Machine-coding of the data would be an expensive procedure due to the size of the SNOP vocabulary. Questions submitted for searching must also be in coded form using AND, OR, and AND NOT logic to express relationships. Searching is performed by means of code comparisons. Output displays may consist of complete autopsy records or of listings of all diagnostic findings that meet the question, showing the case

registry number for each occurrence, and ordered by morphology, topography, etc.

D.W. Crocker [7] suggests a method by which anatomical pathology data could be prepared for automation; however, no mention is made of any algorithm by which one could search the data. The first card of a keypunched autopsy report is termed the Master Card and contains general information pertaining to the case, very similar to the general information division of the Autopsy System described in this thesis. The pathological diagnosis is coded using the SNOP scheme with the suggestion that coding be done manually and each card contain a code followed by the English decoded form. Following the pathological diagnosis section is one card for weights and measurements. On the one card, space is provided for recording the measurements of seventeen particular items; there is no facility to allow the entry of a measurement that is not one of the specified items.

Dr. S.H. Paplanus, Dr. R.H. Shepard, and J.E. Zvargulis [14] present a method for the storage and retrieval of autopsy data that is quite sophisticated; however, like most other systems, it only deals with the final pathological diagnosis. The method employs the use of natural language. A medical secretary keypunches the data onto cards and this is then transferred to tape. Each entry in the final pathological diagnosis is then extracted along with its associated autopsy number and sorted into alphabetical

order. The sorted data is then further manipulated to eliminate duplicate diagnoses. One copy of each diagnosis is retained along with the numbers of all autopsies in which it appears. This information is printed out in the form of a book with page numbers and accompanied by a table of contents for easy reference. When new data is merged with the old, a new book and table of contents is printed. This portion of the system is valuable for guidance in the correct vocabulary usage in formulating a question, as well as serving as a quick reference; however, the cost involved (approximately \$600 for 800 cases) does not seem to be compensated by its usefulness. The system also has a computer searching facility which will search the original tape, i.e. the one onto which card images were written, for either a string of characters or for responses to a question. The authors make no mention of the searching algorithm used other than that it is a standard search program. Output from a search consists of a listing of the stored contents of each autopsy report satisfying the question criteria.

CHAPTER III

FORMULATION OF A QUESTION

In order to retrieve any portion of the autopsy data base as stored on magnetic tape, there must exist some method of conveying a user's desires to the searching phase of the University of Alberta Hospital Autopsy Storage and Retrieval System. This is accomplished by means of the question, or profile, submitted by a user. The searching phase analyzes the question, producing tables, which it in turn uses to perform comparisons against the autopsy data base. If a particular autopsy satisfies all the criteria within a question, it will be included in the results for the search.

A user, or group of users, may submit up to ten questions in a single run, if convenient. Beyond ten queries, additional runs are required.

Every question is composed of two types of cards - the 'question card', which identifies the beginning of a question, and the 'term card', which informs the searching program of the desired searching criteria.

The reference to 'cards' requires clarification. A question may be submitted to the System on punched cards or via the terminal keyboard. The searching phase requires that the input questions be located in a Michigan Terminal System (MTS) file and either method of submission achieves this requirement. Every line within this file contains the

equivalent of a card image, i.e. 80 characters, whether entered by card or terminal, and thus any future reference to cards may be interpreted as either a card or a line.

3.1 The Question Card

Each question card indicates the start of a question, and is composed of four fields, three of which are of relevance to the user: identifier, comment, and output. The following diagram illustrates the location of each field on a punch card.

The identifier field, which must contain the letters QUE, advises the system of the beginning of a new question and that the cards following, until the next QUE is encountered, are to be interpreted as term cards.

The comment field is provided to allow the user to have his output identified, the contents of the field appearing at the top of each page of the final output for the question. In the case where several people have submitted questions to be processed in the same run, it is advisable that the user's name appear in the comment field so that identification of the output is simplified. A good policy to adopt is that each user follow his name with some statement regarding the reason for submitting this query, serving to

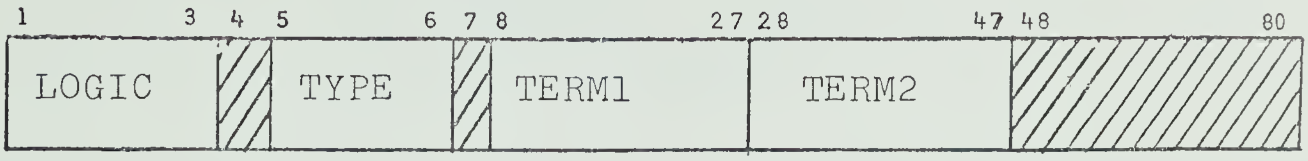
identify output from multiple questions submitted by a single user as well as being a quick reminder to someone looking at output several weeks after it was originally obtained.

A user has the choice of two types of final output and his preference may be expressed by the use of the output format field. If a '1' is sensed in column 80, output will consist of all the information stored for every autopsy fulfilling the requirements of the question. Leaving the output format field blank implies the desire for abbreviated output. The abbreviated form consists of the name and post-mortem number of each case satisfying the requirements of the question. This type of output would be very useful in the event that a user expects many autopsy reports to be output and a complete detailed listing of each is not required, or when a user must examine the anatomical diagrams which are not stored on the tape, or when the year-end statistics must be compiled. The abbreviated output will lead the user to particular cases within the files of the Department of Pathology and probably eliminate what might have been a lengthy manual search.

3.2 The Term Card

By means of the term card, of which there are normally several within a given question, the user has the ability to inform the searching phase of the System as to what his

requirements are. A typical term card contains four usable fields: logic, type, and two term fields.



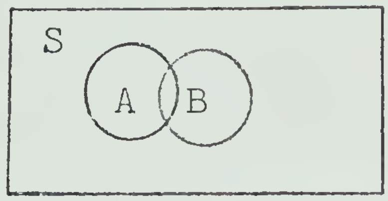
The following sections contain a detailed description of the different options available for each of these fields.

3.2.1 Search Logic

Built into every query language, i.e. a language designed to convey a user's requirements to a searching system, must be some form of logic capability which would allow one to designate relationships amongst terms in the question. The most suitable and probably the easiest for a user to comprehend is Boolean logic. Boolean logic provides an organization of elements adequate for an information system yet is quite simple.

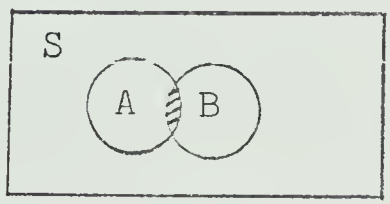
The System allows the use of four types of Boolean connectors: AND, OR, NOT, NOR. In order for these to be better understood some fundamentals of Boolean Algebra will be discussed.

Let S be some set of objects and A and B be subsets of S, i.e. A and B are contained within S. The following Venn diagram might provide a better visualization of the situation.



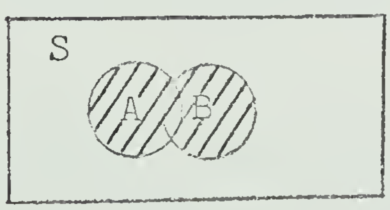
Set S may contain many more subsets; however, only two are being used for the purpose of illustration.

The AND connector, designated by a '·' in Boolean notation, enables one to form a third subset, C, composed of all elements which are in A and in B. The cross-hatched area of the following diagram illustrates this example, i.e. A·B.



A·B

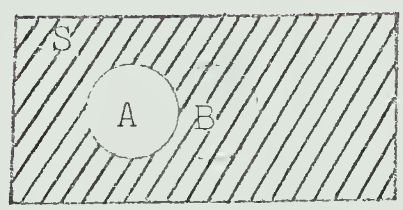
The OR connector, denoted by a '+' sign, creates a new subset satisfying the condition that all its members are in A or in B. The case of A+B is demonstrated in the following diagram.



A+B

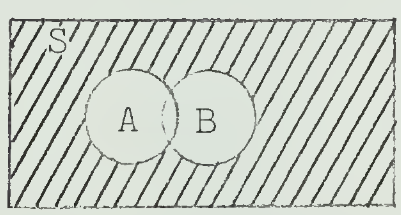
Note that OR is inclusive of elements that are in both A and B, i.e. equivalent of 'AND/OR'.

The NOT operator is normally applied to a single subset. For example NOT A, denoted '$\bar{A}$', refers to all items in class S that are not contained in subset A. From this definition we obtain the following Venn diagram for $\bar{A}$.



$\bar{A}$

The NOR connector is not a single operator, but rather a combination of NOT and AND. Using A, B, and S as previously defined, NOR allows one to generate a subset of S which does not contain any elements from A or from B. The Boolean Algebraic notation for this example is $\bar{A} \cdot \bar{B}$.



$\bar{A} \cdot \bar{B}$

The Autopsy System makes use of the above four types of logical connectors in order to allow a user to specify connections between terms within a question. Even though the usage format is not exactly as described above, the meanings of the operators remain the same. The first difference

stems from the fact that within the System only one term is allowed per card and therefore all desired logical relationships must be between terms on separate cards. A second difference is that every card must contain an entry in the logic field. Deviating from the defined format of the term card, the upcoming example uses only the logic and term fields to demonstrate the above-mentioned differences as applied to the case of A·B.

```
AND  A
AND  B
```

It should be noted that the AND preceding A is required and an error message, indicating that the question was invalidated, will be conveyed to the user if it is omitted.

In the following example an interrogator is interested in obtaining all elements of the data base that contain the terms A or B or C, contain D or E, and it contain neither F nor G nor H and must contain J.

```
AND  A
OR   B
OR   C
AND  D
OR   E
NOT  F
NOR  G
NOR  H
AND  J
```

The only similarity between this example and a question that might be submitted to the System is the usage and placement of the logical connectors.

The logic field of every term card must be filled with one of the above-mentioned options, which must appear left-

justified in columns 1-3. If this format is not adhered to, the question involved will be invalidated.

3.2.2 Search Types

The search TYPE field, located left-justified in columns 5 and 6 of the term card, identifies which part of the autopsy report, e.g. pathological diagnosis, medical history, the System should search to locate the given term. For example, if one were interested in a drug, it would be a waste of time to search the pathological diagnosis division; however, the ability to direct a search to the therapy section, wherein any reference to the drug would be found, would be much more desirable. The System requires that the user direct the search to the relevant portion of the autopsy report. The many options available for the TYPE field gives a user accessibility to almost every portion of an autopsy report. The following sections give a detailed description of the options as well as the required formats for the term fields associated with each.

3.2.2.1 Name

One method of gaining access to a particular case is by searching the data base for the name of the deceased person. This is done by specifying an 'N' in the TYPE field and the desired name in the TERM1 field. Since the System makes literal comparisons, appropriate conventions

must be adhered to in order to ensure that the requested name be in the same format as the name stored in the data base. The System allows a name to have a maximum of sixteen characters including embedded blanks. A suitable convention would be to have the first sixteen characters of the surname, the complete surname followed by a comma followed by part of the first name, or the complete surname, a comma, and the complete first name. If one requires less than the maximum number of allowable characters to relate this information, the remainder of TERM1 should be left blank. If a surname is exactly fifteen characters long, it must still be followed by a comma in the sixteenth position. In all cases the last four positions of TERM1 and the complete TERM2 field must be left blank. The following examples demonstrate several of the existing possibilities.

(1) AND N FEATHERSTONEHAUG

(2) AND N MACDONALDSON,WIL

(3) AND N DOE,JOHN

In the first example only part of the lengthy surname fits into the allotted space; the second case demonstrates a complete surname and part of the first name; the final example contains the complete first and last names with place to spare.

3.2.2.2 Post-Mortem Number

A second method of gaining access to a particular

autopsy report is by specifying the post-mortem number. Facilities have been included to permit one to retrieve all autopsies with a post-mortem number greater than or equal to, less than or equal to, or simply equal to some number specified in the TERM1 field. The main reason for implementing the inequality feature was to simplify the compilation of statistics for the year-end reports.

A post-mortem number consists of 7 characters in the form YY-SSSS, where YY represents the last two digits of the year in which the post-mortem was performed, and SSSS is the sequence number of that post-mortem within year YY. For example, 70-0141 is autopsy number 141 of the year 1970. If one were interested in obtaining all autopsies performed in 1970, one would search for all post-mortem numbers greater than or equal to 70-0000 and less than or equal to 70-9999.

As outlined above, there are three methods of using the post-mortem number to access autopsy reports and corresponding to each is a mnemonic that must be entered in the TYPE field of the term card, i.e. columns 5-6.

PL = post-mortem number less than or equal to
PE = post-mortem number equal to
PG = post-mortem number greater than or equal to

The number that comparisons are to be made against is located in the first 7 locations of TERM1. The remainder of TERM1 as well as the complete TERM2 must be left blank. The term cards associated with the last given example would be:

```
AND PL 70-9999
OR  PG 70-0000
```

3.2.2.3 <u>Age</u>

The Autopsy System allows a user to search the age field of the autopsy report and retrieve all cases satisfying the criterion specified in TERM1. As in searching on post-mortem number, this capability has been extended to include inequality searches. This allows one to restrict output to only those cases that satisfy a given inequality, where once again the comparison is done against the age given in TERM1. The three possibilities for placement in the TYPE field are:

```
AL = age less than or equal to
AE = age equal to
AG = age greater than or equal to
```

In order to allow for ages greater than 99, the age field must be three characters long. The age of interest is placed in the first three positions of TERM1, leaving the remainder of TERM1 plus all of TERM2 blank. The specified age must be right-justified within these three positions and all blanks at the left filled by zeros. For example,

```
age  102  should appear in TERM1 as  102
age   81  should appear in TERM1 as  081
age    6  should appear in TERM1 as  006
age    0  should appear in TERM1 as  000.
```

Since the System only allows for age specifications in years, use in the normal manner would preclude separation of stillbirths from deaths in the first year. To overcome this, age could be interpreted generally as 'age next birthday', allowing stillbirths to be specified as age 0.

3.2.2.4 Sex

In order to search for a particular sex, one would place an 'S', left-justified, in the TYPE field, and either F (female) or M (male) in the first position of TERM1. Both of the following examples yield the same net result of retrieving all cases wherein the sex is recorded as female.

(1) AND S F

(2) NOT S M

3.2.2.5 Race

The term card requirements for searching the racial origin entry in the autopsy report is very similar in nature to the search for the sex. The mnemonic to be placed in the TYPE field is 'R' and one of the following possibilities must occur in the TERM1 field:

C = Caucasian
N = Negro
O = Oriental

Once again the user is cautioned that the chosen option must be placed in the first column of the TERM1 field, with the

remainder of TERM1 as well as all of TERM2 left empty.

3.2.2.6 Clinical Diagnosis

In most post-mortem reports there is some short statement giving the clinical diagnosis as to the cause of death. This statement is recorded word for word, preferably without such words as 'a', 'an', 'the', etc., in the autopsy data base. In order to be able to search on any term that may appear in this diagnosis, one would enter 'CD' in the TYPE field and follow this with the desired word in the TERM1 field. If one wishes to do a search on more than one word within the clinical diagnosis, a separate term card is necessary for each word with appropriate logic forming the desired link between the words. The maximum length of any word to be searched for is the 20 characters of TERM1. All words of longer length must be truncated; however, there should be no fear of not obtaining a match between the question term and the term in the data base as all terms within the latter are also truncated to 20 characters. Any requested term of length less than 20 must be left-justified in TERM1.

The following extract from a possible question submitted to the System exemplifies the way one could use this searching capability.

```
AND CD CANCER
OR  CD CARCINOMA
AND CD BREAST
```

It should be noted that AND and NOT, which are of equal precedence, have a higher precedence than OR and NOR, which are also of equal precedence. Using Boolean Algebraic notation and brackets to show precedence, this example would be expressed as follows:

$$(CANCER+CARCINOMA)\cdot BREAST$$

The user is interested in all cases where the clinical diagnosis indicated breast cancer. The words cancer and carcinoma are used synonymously and thus the user has used OR logic to ensure that this portion of the question return affirmatively if either of these terms is encountered in the clinical diagnosis.

3.2.2.7 Pathological Diagnosis

A major portion of any autopsy report is devoted to stating the pathological diagnoses, i.e. the detailed findings of the post-mortem examination. A user of the Autopsy System has the ability to direct the System to perform any one of three types of searches on the pathological diagnosis division. Each requires the mnemonic 'PA' in the TYPE field; differentiation is by use of the TERM1 and TERM2 fields.

The TERM1 field is reserved for the name of any part of the human body and TERM2 embodies any requested descriptor. The adjectives 'left' and 'right', as used to differentiate between pairs of organs, must appear preceding the organ

name in TERM1 and assume the form LT or RT. For example, the left lung would be specified at LT LUNG. The three permissible methods for searching the pathological diagnosis portion are: (1) organ, (2) organ plus descriptor, (3) descriptor only.

In the first instance, the desired part of the body is placed in TERM1 and TERM2 is unused. The contents of TERM1 must be left-justified with words of length greater than 20 characters truncated. This method of searching allows one to retrieve all cases wherein the mentioned body part appears in the pathological diagnosis. If a user were doing a study on diseases of the lung, the following might be used as part of a question.

```
AND PA LUNGS
OR  PA LT LUNG
OR  PA RT LUNG
```

This example will assure retrieval of all cases with any reference to lungs, right lung, or left lung in the pathological diagnosis.

The second search possibility allows one to look for any specified portion of the human anatomy as well as some term used to describe the given body part. The body part appears in TERM1 and the descriptor in TERM2, with the left-justification and truncation rules applying to each. Extending the previously given example one could search for all mentions of lung congestion in the following manner.

```
AND PA LUNGS          CONGESTION
OR  PA RT LUNG        CONGESTION
OR  PA LT LUNG        CONGESTION
```

The final method allows searching for a descriptor, as for example carcinoma, irrespective as to which organ it may apply to. The TERM1 field of the term card is left blank and the desired descriptor placed in TERM2. The following example searches the pathological diagnosis for pneumonitis of the right lung as well as any occurrence of cancer.

```
AND PA RT LUNG        PNEUMONITIS
AND PA                CANCER
OR  PA                CARCINOMA
```

3.2.2.8 <u>Measures</u>

Nearly every post-mortem report includes measures of some body organs or liquids. This information is recorded in the measures section of the autopsy protocol and is available for searching by any user of the System. It would be quite rare to want to search for an exact numerical measure and therefore inequality searches have been incorporated. In order to search the measures division, one of the following mnemonics must be placed in the TYPE field.

ML = measure less than or equal to
ME = measure equal to
MG = measure greater than or equal to

The organ or fluid name, which must abide by the left-

justification and truncation conventions, is placed in the TERM1 field. The measure against which the search is to be performed must be left-justified in TERM2. In order to facilitate use of the System, there are no options regarding units of measurement. Length, weight, and volume are each to be recorded in one standard unit. Centimetres, grams, and millilitres have been presumed as these units and the storage capacities set accordingly. The actual units are never quoted; they must be determined by context.

The numerical measure must be expressed in the form NNNNNN.NNN, i.e. in ten characters, and placed in the first ten positions of TERM2. The six digits prior to the decimal point and three thereafter should suffice every measurement requirement, from the smallest organ size to the complete body weight. Every one of the allowable ten positions must be filled, even if it be with a redundant leading or trailing zero. The following examples illustrate the forementioned conventions.

(1) 20.1 must be expressed as 000020.100
(2) 1987 must be expressed as 001987.000
(3) .04 must be expressed as 000000.040

If one does not adhere to this format, the obtained results will most probably not conform to the user's intentions.

The following term card prototypes show the varied ways in which one may use the capability to search the measures portion of the autopsy report.

(1)	AND ML RT KIDNEY OR MG RT KIDNEY	000065.000 000120.000
(2)	AND MG MITRAL VALVE	000009.000
(3)	AND ME BRAIN	001016.000

In the first example the understood unit of measurement is grams and the user is interested in all right kidneys with a weight of ≤65 grams or ≥120 grams. It should be noted that LT and RT are used in the same manner as in the pathological diagnosis. The remaining examples are straightforward with the units being centimetres and grams respectively.

3.2.2.9 Drugs and Other Therapy

An extremely important and valuable portion of any post-mortem report is the section describing the therapy administered to a patient. With a knowledge of the cause of death, the therapy administered, and a substantial data base one may research the conditions which cause a particular type of therapy to have a beneficial effect in some cases of a disease and an adverse one in others. One also has the ability to study the after-effects from the administration of a drug. This information could help doctors in determining what type of therapy to administer and if an extra measure of caution should be taken in its administration.

The System contains eight searchable subdivisions of each entry within the therapy section - therapy name, dosage,

rate, route, start date, end date, reason administered, and sequel. The following paragraphs outline the rules and conventions governing the usage of each one.

In order to search for a therapy type by name, one must place the mnemonic 'TH' in the TYPE field followed by the name in the TERM1 field. The contents of TERM1 must be left-justified, truncated to 20 characters, and followed by a blank TERM2 field.

AND TH THIOTEPA

As in the specifications for body measures, all dosage representations are expressed in standard units. These units are either millilitres or grams, as appropriate, and no mention of them need be made in a search. The System allows one to perform inequality searches for the therapy dosage by using any of the following for the search type.

DL = therapy dosage less than or equal to
DE = therapy dosage equal to
DG = therapy dosage greater than or equal to

In order to assure that a dosage applies to the therapy of interest, the therapy name must appear in TERM1 in the same format as described above. The dosage is specified in TERM2 in the same format as used for the measures, i.e. in the first ten columns of TERM2 and consisting of six integral positions followed by a decimal point and then three decimal places. The user is reminded that within this ten character field all redundant leading and trailing positions must be

filled with zeros.

```
AND DG THIOTEPA          000001.500
AND DL THIOTEPA          000003.250
```

This example requests all cases wherein a dosage, D, of the drug Thiotepa was administered, such that $1.5 \le D \le 3.25$.

In order to be able to compute the total therapy dosage given a patient, one must have available the rate of administration, dosage, starting date, and the ending date. By placing 'RA' in the TYPE field, the therapy name in TERM1, and the rate code in TERM2 one may search the rate of administration subdivision. The contents of TERM2 must be left-justified and conform to the normal coding scheme used by the Department of Pathology at the University of Alberta Hospital.

```
HS  = at bedtime
QID = 4 times per day
TID = 3 times per day
BID = 2 times per day
QxH = every x hours (x=1,2.....etc.)
PRN = as required
PC  = after meals
AC  = before meals
QD  = once a day
```

This scheme can be expanded as necessary, providing that each entry does not exceed 8 characters in length. The searching phase is dependent on the contents of the autopsy data base and therefore all one need do to make additions to the coding scheme is enter the new code in the correct portion of the autopsy report in all cases where it is applicable

and notify other users of its availability.

```
AND RA THIOTEPA        QID
OR  RA THIOTEPA        TID
```

The route of administration conveys information regarding the method by which the mentioned therapy type was applied to the patient. By specifying 'RT' in the TYPE field, one has the ability to search for any route of administration that may be of interest. In order to ensure that the given route applies to the therapy of interest, the therapy name must be given in TERM1. The route code, placed left-justified in TERM2, can be any word of up to eight letters but, to allow ready comparisons, should be from a standard code, as for example

IV = intravenous
IM = intramuscular
ID = intradermal
IT = intrathecal
SC = subcutaneous
PO = orally
PR = per rectum
PV = per vaginum
IA = intra-arterial
IC = intracardiac
IJ = intraarticular
IP = intraperitoneal

In the following example the user, as a partial requirement for his query, desires all cases in which a patient was administered the drug Gentomycin intravenously.

```
AND RT GENTOMYCIN        IV
```

Aside from being required to obtain the total dosage

of therapy administered to a patient, the therapy start and end dates enables one to retrieve cases of patients who were treated with a drug that was later discovered to be part of a faulty batch. In order to have this capability, inequality searches on both the starting and ending dates are a necessity. The following are the permissible mnemonics for the TYPE field of the term card.

SL = therapy start date less than or equal to
SE = therapy start date equal to
SG = therapy start date greater than or equal to
EL = therapy end date less than or equal to
EE = therapy end date equal to
EG = therapy end date greater than or equal to

As is the case for all subdivisions of the therapy portion of the autopsy report, the therapy name must appear in TERM1. The date, specified in TERM2, is in the form YYYY/MM/DD, where YYYY are four digits corresponding to the year (e.g. 1971), MM is a two digit numerical representation of the month of the year (e.g. March = 03), and DD is the day of the month. Every date requested must completely fill the first ten positions of the TERM2 field. In many cases, the exact date is not known - possibly only the year, or the year and month are known. In order to alleviate this problem, the month and the day of the month are allowed to assume a value of 00. If only the year and month are known, these fields should be filled in appropriately and DD set to 00. If only the year is available, the month and day should be 00. With the aid of several examples, any existing

uncertainties should be clarified.

(1) AND SL THIOTEPA 1969/01/20

(2) AND SL THIOTEPA 1970/02/99

(3) AND SL THIOTEPA 1970/00/00

In the first example, the searching phase of the System must find all cases wherein Thiotepa was administered on or prior to January 20, 1969; in (2) the start date must have been prior to March 1970; the final example will retrieve any administration of Thiotepa prior to 1970, as well as those cases where it was administered during 1970 and the exact month is unknown, i.e. the date appearing on the autopsy tape is in the same form as that in TERM2.

Associated with any order for the administration of therapy must be a reason for that order. In all instances where this information is thought to be of value and retained, it is recorded on the autopsy report and subsequently made available to the System user. Some research interests may require a knowledge of all the drugs administered to combat a particular illness, as for example all therapy used in the treatment of jaundice; whereas other interests may desire all instances of a particular therapy type being administered for a particular reason, as for example the usage of Pilocarpine drops as an aid to curing glaucoma. By placing 'TR' in the TYPE field and appropriately using TERM1 and TERM2, the Autopsy System allows one to

search for either of the mentioned possibilities. The TERM1 field is reserved for the therapy name and should be left blank if one is interested in obtaining the names of all therapies associated with the reason for administration mentioned in TERM2. The TERM2 entry must be left-justified, truncated to 16 characters, and may only consist of a single word. If the desired reason for administration is several words long, a new term card is required for each word. For example, pleural effusion would require the use of AND logic to make the appropriate link and ensure the presence of both words. By using both term fields one may obtain all relationships between the administration of the therapy specified in TERM1 and the TERM2 reason for administration.

```
(1)  AND TR THIOTEPA          PLEURAL
     AND TR THIOTEPA          EFFUSION

(2)  AND TR                   GLAUCOMA
```

In example 1 the user is requesting all cases in which Thiotepa was used to counteract pleural effusion; the next example requires all cases that suffered from glaucoma and had some form of therapy used in an attempt to overcome it. By searching for a therapy name one would automatically obtain all the reasons why this particular form of therapy was administered, and thus a search of this nature was not incorporated into the reason for administration capabilities.

Searching the therapy sequel, i.e. what happened after the administration of the therapy, is nearly the same as

searching the reason for administration. The usage of the TERM1 and TERM2 fields remains unchanged. The only difference is that the TYPE field mnemonic must be 'TS'.

(1)	AND TS THIOTEPA	LEUKOPENIA
(2)	AND TS	JAUNDICE

In the first case all occurrences of leukopenia resulting after treatment with Thiotepa will be retrieved; in (2) the results will include cases in which jaundice developed after any type of therapy treatment.

3.2.2.10 Medical History

The last section of each post-mortem report usually contains a medical history of the deceased person, normally containing many details for the period immediately preceding death.

Each entry in the medical history division consists of a date followed by a description of what happened on that date. In order to search this portion of the autopsy report, 'HT' must appear in the TYPE field. By appropriately making use of the two term fields, it is possible to do any of the following: (1) search for a date; (2) search for a date and happening; (3) search for a happening.

The TERM1 field is reserved for a date and TERM2 contains a word, left-justified and truncated to 20 characters, that one would like to search for. The date must be in the

same format as described for the therapy start and end dates. Unknown and partially known dates must also conform to the previously defined conventions. If a user should desire to search for a string of descriptors that might appear together, as for example Bell's palsy, the terms must appear on separate term cards and connected by AND's, except for the position words 'left' or 'right' in conjunction with a body organ name which must be specified as a single term, e.g. LT BREAST.

By filling in TERM1 and leaving TERM2 blank, one would be requesting a search on a particular date and nothing else. The requested date must be exact as no inequality searches are permitted on this portion of the autopsy report. The use of both term fields requests the searching phase to look for the given date and descriptor. Leaving TERM1 blank and supplying information in TERM2 directs the search to disregard all dates and only look for the supplied descriptor. The following three examples demonstrate the capabilities of the history search as described above.

(1)	AND HT	1971/02/25	
(2)	AND HT	1970/03/10	SINUS
(3)	AND HT		CARCINOMA

3.3 <u>An Example</u>

The following example is given to summarize the above and demonstrate the concept of a complete question as it

would be formulated for submission to the University of Alberta Hospital Autopsy Storage and Retrieval System.

```
QUE     MARVIN BARUDE. SAMPLE QUESTION
AND N   SMITH,MARY
OR  N   CASE,NO4
AND PG  70-0000
NOT PE  70-0050
AND S   F
AND R   C
OR  R   N
AND AG  080
AND CD  CANCER
OR  CD  MASTECTOMY
AND TH  THIOTEPA
OR  TH  PILOCARPINE DROPS
AND SG  PILOCARPINE DROPS      1953/00/00
AND EL  PILOCARPINE DROPS      1971/00/00
AND RT  THIOTEPA               PO
AND TR  THIOTEPA               PLEURAL
AND TR  THIOTEPA               EFFUSION
AND TR                         GLAUCOMA
AND TS  THIOTEPA               LEUKOPENIA
OR  TS                         JAUNDICE
AND MG  BODY LENGTH            000100.000
AND PA  LT LUNG                THROMBOEMBOLUS
OR  PA  VEINS                  THROMBOSIS
AND PA                         CARCINOMA
AND HT  1961/10/27             CARCINOMA
AND HT                         SINUS
OR  HT  1969/10/30
```

The question-batching facility may be made use of by immediately following the first question with up to nine additional ones. By attaching the following question to the previously given sample, we will obtain a two-question batch.

```
QUE     MARVIN BRAUDE. SAMPLE QUESTION 2
AND R   C
AND AL  090
NOT PL  70-0000
NOR PG  70-9999
```

Note that each question is entirely independent; anything in common between the two has to be explicitly repeated.

It was necessary, in developing the System, to provide a limit to the number of AND and NOT statements in any question. This has been set at 25, i.e. the number of AND plus NOT cards may not be greater than 25. There is no limit as to the number of OR's or NOR's, other than that the total number of cards within a batch of questions may not exceed 100.

3.4 Summary

The full range of TYPE field options are herein summarized, for quick reference.

N	=	name
PL	=	post-mortem number less than or equal to
PE	=	post-mortem number equal to
PG	=	post-mortem number greater than or equal to
AL	=	age less than or equal to
AE	=	age equal to
AG	=	age greater than or equal to
S	=	sex
R	=	race
CD	=	clinical diagnosis
PA	=	pathological diagnosis
ML	=	measure less than or equal to
ME	=	measure equal to
MG	=	measure greater than or equal to
TH	=	therapy name
DL	=	therapy dosage less than or equal to
DE	=	therapy dosage equal to
DG	=	therapy dosage greater than or equal to
RA	=	therapy administration rate
RT	=	therapy administration route
SL	=	therapy start date less than or equal to
SE	=	therapy start date equal to
SG	=	therapy start date greater than or equal to
EL	=	therapy end date less than or equal to

EE	=	therapy end date equal to
EG	=	therapy end date greater than or equal to
TR	=	reason for therapy administration
TS	=	sequel to therapy administration
HT	=	medical history

CHAPTER IV

QUESTION PROCESSING

4.1 Creating the Question File

Since considerable thought is usually required to design questions which will correctly convey the user's requirements to the System, it is advisable to begin by outlining the desired query on paper. Let us assume that a user has formulated what he believes to be a logically and syntactically correct question and must now proceed to enter it into the MTS file QUESTIONS as the first step toward obtaining results. At this point two options are available: (1) punch the questions on cards; (2) enter the questions via the terminal keyboard.

Once a user is signed on, i.e. has completed the preliminary procedure of identifying himself to MTS, whether he be in terminal mode or batch mode, the following series of instructions will create the required file for the questions.

```
$EMPTY QUESTIONS
$GET QUESTIONS
$NUMBER
            }  the questions
$UNNUMBER
```

If batch mode is being employed, the questions, as punched on cards, are placed between the $NUMBER and $UNNUMBER

command cards. In on-line mode, after transmitting the $NUMBER command, the system will respond with '1' as the line number. At this point the user should type the first statement of the question batch, considering the position where the carriage stops as position 1 on a card, and then depress the RETURN key, after which '2' will appear as the next line number. This process continues until the complete batch of questions has been entered into the file. Then, to signify completion to the computer, the user should type $UNNUMBER following the line number.

Upon completion of the preceding operations, the user is once again confronted with two possible paths - call the file editor to correct any errors in the file QUESTIONS or call the analyzing and searching phases of the System.

4.2 Error Correction Within a Question

Errors during entry of questions must be expected for even the most cautious user. Regardless of how major an error is, correction is always possible prior to calling the question analyzing and data base searching portions of the retrieval system.

The most likely place for errors to be discovered is while seated at the keypunch or terminal entering lines into the question file. Correcting errors within a card deck simply requires the replacement of erroneous cards or the insertion/deletion of additional cards.

If an error is discovered during an on-line terminal session, correction is a little more complicated. If the error is in the line presently being entered (probably a typing error), backspacing to the position of the error and retyping everything from there on will make the proper adjustments. Alternatively, one may enter an understroke followed by a carriage return, which will result in the complete line being deleted and MTS returning with the same line number on a new physical line, where the user may now recommence.

If the fault is in a previously transmitted line, it is suggested that the user continue entering the remainder of the questions and only make corrections, in a manner to be described in future paragraphs, after the $UNNUMBER command.

In many instances errors may be discovered after output has been received; however, the ability to make corrections and resubmit the questions still exists. The Autopsy System has the ability to detect syntax errors, e.g. placement of the TYPE field in the wrong card columns. Resulting from this type of error will be the System generated diagnostic ***QUESTION INVALID - ERROR IN NEXT CARD***, which appears in the output listing of the question preceding the erred line. Unexpected results may lead a user to reexamine his question and possibly discover previously unnoticed errors.

In most instances, rather than emptying QUESTIONS and recommencing, one may make use of the MTS file editing capa-

bilities. The MTS File Editor is a program which allows one to perform many varying editing operations on an existing file. It is suggested that the user refer to the University of Alberta Computing Center publication 'File Editor' in order to familiarize himself fully with the capabilities and usage of the editing facility. Of particular pertinence to the Autopsy System are the editing commands

```
COLUMN  = set the column range for editing operations
DELETE  = to delete a line
INSERT  = to insert a line
PRINT   = to print a line
REPLACE = replace the contents of a line with the
          given string
OVERLAY = to overlay a line with a given string
STOP    = to stop execution of the editor
```

Prior to being able to use these commands, the File Editor, which is stored in the public file *EDIT, must be invoked. This is done by entering the command

```
$RUN *EDIT
```

Upon MTS's acceptance of this command, the user is prompted for the name of the file to be edited.

```
:ENTER FILE NAME:
:
```

At this point the user should enter QUESTIONS, hit the RETURN key, and await the appearance of the next colon before attempting to use any of the available commands.

Editing operations may be performed in either batch or terminal mode. Because of its dialogue facilities and the

greater ease in avoiding further errors, the terminal mode is recommended. For instance, upon acceptance and completion of an operation, the editor displays the new contents of a line or portion of the file. In terminal mode this is immediate and isolated, improving user attention and allowing immediate correction if still in error; in batch mode the user sees the output only after it is all assembled and printed.

In the worst possible case, one must order the file QUESTIONS emptied and re-enter the questions. The MTS commands required to do this are:

```
$EMPTY QUESTIONS
$GET QUESTIONS
$NUMBER
                    correct version
$UNNUMBER
```

4.3 The Searching Phase

At this stage the supposedly error-free questions reside in the QUESTIONS file and the user may now proceed to set the analyzing and searching phases to work. These routines analyze the submitted questions and perform logical comparisons against every autopsy in the data base. The final results are output either at the terminal or on the line printer, depending on which mode of operation is being

employed.

In most instances a user, having established his questions within the System, will proceed immediately to have the appropriate analyses effected. Alternatively, the QUESTION file may be created and left dormant for a time, though this requires that no one else uses the System and overwrites the questions in the meantime. In the normal case, the user is already signed on and may continue in one of the manners described hereafter; in the second case, he must first sign on and then continue.

The three means by which one may submit the QUESTIONS file to the System for processing are: terminal, batch, or batch from the terminal. The advantages and disadvantages of each will be discussed in a later section; however, an explanation of how to use each follows. It is assumed that the user is signed on prior to attempting to proceed any further.

In the on-line mode, the user at the terminal initiates the search and waits for the results to come back to him. The user need only give the system the following command and everything else will be done automatically for him, including the sign off.

```
$SOURCE SEARPROC
```

In batch mode this same command punched on a card, will perform the equivalent operations. The only difference

between the two is that the output will now come out on the line printer rather than on the terminal.

Normally a person connected in the on-line mode will perform all input to the system and receive all output from the system at the terminal, i.e. he has no access to the card reader or the line printer. The terminal prints at a small fraction of the speed of the line printer and if a considerable amount of output is expected, it would be both costly and protracted to use the terminal. MTS has a facility where one can submit a batch job from the terminal, i.e. the system commands are given from the terminal but are accepted as if they came in on cards for batch mode and any printed output is directed to the line printer. The following sequence of commands will enable a user of the Autopsy System to make use of this capability

```
$RUN *BATCH
$SIGNON   etc.
password
$SOURCE SEARPROC
$ENDFILE
$SIGNOFF
```

The entry of *BATCH following $RUN has prompted the terming of this mode of computer use as '*BATCH' (="star batch"). It must be emphasized that even though a user is signed on prior to submitting the above commands, he must still sign on again in the indicated place. After the $ENDFILE the MTS system will return a receipt number to the user, which should be noted and given to the clerk at the Computing

Center's input-output office for retrieval of the output. This method does not automatically sign off a user from the terminal, as do the previous two methods, and therefore an explicit $SIGNOFF command must be transmitted. It is imperative that the sign off command not be forgotten since the *BATCH job will not be able to be initiated if the user identification number is busy. The user is warned against entering new questions into the QUESTIONS file prior to the *BATCH job having run to completion.

After signing on, one may check the status of any batch job in the system, i.e. its position on the execute or print queue, if presently executing or printing, or if it is not found (completed). This information will help the user to determine whether he can submit a new batch of questions, or usefully make alterations to questions already in the file. In order to gain the information regarding the status of a job, the following command must be relayed to MTS.

$RUN *HBQ PAR=receipt number

The receipt number is that number generated by the system upon acceptance of a *BATCH job or the number on the receipt card given to a user when he submits a card deck personally. If a pickup and delivery service is employed, the receipt card will not be seen by the user.

4.3.1 The SEARPROC File

In the design stage of the Autopsy System it was taken into consideration that the majority of users would be medical staff with very little acquaintance with computing science. To meet this constraint the System was designed so as to minimize the amount of effort and computer knowledge required to obtain results.

In all three methods of initiating the search, the command $SOURCE SEARPROC is made use of. MTS provides the ability for one to place a string of MTS commands in a file and subsequently direct MTS to this file for its instructions. Usually the terminal or card reader, depending on the mode of operation being employed, is the source for input commands; however, the $SOURCE command allows one to change the command source.

In our case the pertinent instruction string for running the search is stored in the file SEARPROC, making usage of the system much simpler. The following is a listing of the contents of SEARPROC with a brief description given thereafter.

```
1.  $RUN *MOUNT PAR=0034 9TP *TAPE* VOL=T00034 SIZE=7280
          LRECL=80 FMT=FB RING=OUT
2.  $CREATE -SORTIN SIZE=100P
3.  $RUN INQUIRY 1=QUESTIONS 5=*TAPE* 6=-SORTIN
4.  $CREATE -SORTOUT SIZE=100P
5.  $RUN *SORT
6.  SORT=CH;A;121;9;CH;A;130;4
7.  INPUT=-SORTIN;F;133;133
8.  OUTPUT=-SORTOUT;F;133;133 MNR=50000
9.  $RUN *EDIT SPRINT=*DUMMY* GUSER=*SOURCE*
```

```
10.  -SORTOUT
11.  REGION /A 1 *L
12.  COL 121 133
13.  B /A'           '
14.  STOP
15.  $COPY -SORTOUT TO *SINK*
16.  $SIG
```

At 1. the computer operator is requested to mount the tape containing the autopsy data base on a 9-track tape drive with the ring out, i.e. allowing reading from the tape but not writing on it. The additional information describes to MTS the format of the data on the tape. Step 2 creates a file to hold the output of the search. In Step 3 the object module of the searching program is run, with any reference to hardware unit 1 implying the QUESTIONS file, unit 5 the autopsy tape, and unit 6 the output file. The next step creates a file in anticipation of output from the sort stage which follows. Steps 5-8 call the sort package, describing the manner in which the sort is to be performed, where to find the input, and where to place the output; sorting is required to arrange the final output such that all answers follow their corresponding questions. Upon completion of the sort phase, the file editor is called in Steps 9-14. In previous steps, information required for proper processing was appended to the output; however, this information is extraneous in the final printed report. The file editor prepares -SORTOUT for final output to the printer or terminal by removing the extra information. The final step, preceding the sign off, is to copy the sorted and edited

file -SORTOUT to the line printer or terminal, depending on the mode of operation being used.

4.4 The Search Output

As previously mentioned, a user has the choice of two forms of output - full or abbreviated. Regardless of the output format desired, each question is re-stated in the output immediately prior to its answer. Where several questions are batched, the form becomes: question, answer, question, answer, etc. The question listing, along with diagnostics, appears on a page by itself and is headed by the title SEARCH QUESTION. Following the question listing, and prior to the answer, a line with a date informs the user of the last time any changes were made to any part of the autopsy data base. The answer to each question starts on a new page.

4.4.1 Full Output

In the case of full output the pages within an answer are consecutively numbered from 1. If the answer contains more than one autopsy report, every such report begins on a new page with the page numbering sequence uninterrupted. At the top of every page appears the heading THE UNIVERSITY OF ALBERTA HOSPITAL AUTOPSY SERVICE, followed on the next line by the identifying comment extracted from the QUE card and the page number. The final line of the header contains

the name of the deceased person, the post-mortem number, and the version date, i.e. the last date that any changes were made to this particular autopsy record. Because of the likelihood of spelling mistakes and other errors in records being noticed during System use, facilities exist to enable alteration of records. The version date must be changed whenever any such changes are made.

With full output, the computer-generated report is substantially a verbatim copy of the autopsy record as entered (and subsequently amended). Following the general information, sectional headings are pathological diagnosis, measures, drugs and other therapy, and medical history.

The general information section contains details such as age, sex, race, University Hospital patient identification number, birth date, date of admission to hospital, date of death, date of autopsy, completeness of autopsy, resident pathologist, staff pathologist, service doctor, and finally the clinical diagnosis.

The pathological diagnosis division is formatted such that every line has four fields: body part, condition, condition, and condition. The condition fields contain the actual pathological findings pertaining to the anatomical part at the beginning of the line. If less than three condition fields are required, the remaining ones are left blank. If any body part should require more than three condition fields the continuation is on the following line

with the body part field left blank.

Each line of output under the measures division contains: item measured, measure, item measured, measure, i.e. every line contains two items of information. No units of measurement are displayed, but rather one should be able to determine if centimetres, grams, or millilitres is implied from the context.

For every type of therapy administered to a patient, three lines of output will result with a blank line separating each type. The first line contains the therapy name, the dosage administered, the rate of administration, route of administration, starting date, and ending date. As in the case of measures, the units for the dosage do not appear but are implied from the type of therapy. The next line contains the reason for administering the therapy; the last line shows any sequel to the therapy. Any one of the above-mentioned fields may be left blank, implying that there is either no information or that the information is of no interest.

Within the final portion of the autopsy report - the patient's medical history - every line has the format: date, sequence number, finding, finding, and finding. The format of the date field, as well as exception handling, was detailed in Chapter III in the section on search types and the reader may refer back to refresh his memory. If a line has the date field blank, the date of the preceding line is to

be assumed. The sequence number allows for the correct ordering of events within a particular date. The next three fields, i.e. the finding fields, describe what occurred on the given date.

Each of the previously described divisions is slightly indented under the appropriate heading. Enclosed in brackets on the heading line are the names of the fields in the order in which they occur within the division. Most autopsy reports will require more than one page of output and if a page break should occur in the middle of a division, the next page will contain the division heading followed by the word CONTINUED immediately after the normal page header. The output from a terminal and the line printer differ slightly. At the terminal, the output page breaks may not occur at a physical page break since, unlike the line printer, the computer has no way of knowing the positioning of the paper. An output page break at the terminal can be recognized by a skip of approximately seven lines and following this the standard page header.

4.4.2 Abbreviated Output

With abbreviated output format only the patient's name and post-mortem number are given for each cited case. The page header takes on the same form as for the full output, with the exception of the final line which contains the heading ABBREVIATED RESULTS in place of the case name,

post-mortem number, and version date. The page numbers occur in the same place and follow the same scheme. However, only one line is devoted to each case retrieved, with new pages being started only as necessary. The abbreviated output can be used as a means of counting cases that meet certain criteria, or for identifying cases for which reference to orthodox records is seen as essential.

4.5 Batch Mode Versus On-Line Mode

Under the MTS operating system one has the choice of using batch mode or terminal mode. In batch mode, input to the computer is by cards which are fed through the card reader; output is normally obtained from the line printer. In terminal mode, all input and output is performed through the terminal. A person using this mode of operation has no access to the card reader or line printer. Within the terminal mode there exists an MTS facility to submit a batch job. This facility termed *BATCH, allows the creation of a batch job from a terminal and and thus provides access to the line printer.

A factor which must be taken into serious consideration is the recent innovation of charging for computing time and facilities. The charging algorithm presently employed allows for the specification of three levels of priority - low, normal, and high. If low priority is specified, the job is placed in hold status and only processed on a first-

in-first-out basis if there are no higher priority jobs in the system. The resulting charge is 70% of the normal charge. In normal priority, the job or sign on card is scanned for the amount of time and number of output pages requested and a further priority assigned to determine where this job should be placed in the processing queue. High priority jobs run on a first-in-first-out basis prior to any normal or low priority job, and the charge is 1.3 times the normal cost. Terminal mode is considered as high priority and is charged accordingly, as well as having an additional charge for connect time.

In trying to determine which mode of operation would be best suited for the different phases of the Autopsy System, the form of output and the access to the Computing Center for batch processing has also to be taken into consideration. As previously described, the output from a terminal is not as appealing as that from the line printer due to the page skipping procedure used. Most users of the System will be members of the Department of Pathology and since that department is located at some distance from the Computing Center, an attempt should be made to avoid numerous cross campus marches. The University provides a pickup and delivery service between the Clinical Sciences Building and the Computing Center twice a day and whenever possible this method of transfer should be used.

After taking all of the above facts into consideration,

the following modes of operation are recommended to the user.

<u>When creating the QUESTIONS file</u> it would be advantageous to use terminal mode since the discovery of an error is easily rectified at the terminal. This method would probably save what would have been a bad run and this in turn will save the user money.

<u>All corrections to QUESTIONS</u> should be done at the terminal for the same reason as mentioned above and also because of the immediate echo facility, whether the job was created via terminal or cards.

<u>Running of the search</u> should be done via *BATCH, to save the trouble of sending over cards to the Computing Center and then having to pickup the output later, unless there is an extreme urgency. The output can be collected from the Center or directed via the pickup service to the Clinical Sciences Building. In cases of urgency, when immediate output is required, terminal mode may be used for all phases of the System.

Additional phases of the Autopsy System, namely those dealing with maintenance of the data base, will be discussed in the next chapter and recommendations as to the mode of operation will be made at that point.

CHAPTER V

THE AUTOPSY DATA BASE

The data base of the Autopsy System consists of reports of autopsies performed by the Department of Pathology at the University of Alberta Hospital. The average autopsy report requires approximately 125 punched cards to record all the information required by the System. Owing to the great number of reports on file in the Department of Pathology, as well as continuous additions to this file, the means of storage made use of must be permanent, economical, and have the ability to hold a large amount of data. There are two types of storage devices available to a computer user at the University of Alberta - disk and tape. The cost of the physical medium is comparable for the two types. However, the operational arrangements at the Center allow only for continuous availability of disk-stored information - with a concomitant cost for the control and drive units that are thereby tied up. Tape, on the other hand, can be kept off-line and mounted on the computer only when required by the Autopsy System. In some instances processing time is greatly reduced if the data is stored on disk and the cost of storage can be compensated for by the saving in processing time; however, since all access to the autopsy data base is sequential, this factor has little effect in our case. Tape was therefore the medium of choice.

The data on the tape is stored as card images, i.e.

each record contains 80 characters of information; it is blocked at 91 records per block. Assuming that a 2400-foot tape is being used with a recording density of 800 characters per inch (bpi) and that the average autopsy report requires 125 records, one reel of tape will hold approximately 2000 autopsy reports. The University of Alberta Hospital performs approximately 500 autopsies per year, from which it follows that one tape is sufficient for four year's work.

The present implementation of MTS creates one problem for users of the Autopsy System in that it does not have the ability to handle multi-reel files. A single tape should be sufficient for immediate requirements; however, it is inevitable that the autopsy data base will outgrow this limitation sincc the initial updating procedure will consist of storing newly completed autopsy reports as well as going backward in time. The Computing Center staff is aware of this problem and has placed it on a list of projects to be investigated when time permits; however, until such time as a solution is found, the user of the Autopsy System will be forced to run the complete System once for each tape in the data base, with the appropriate tape mount commands altered. Keeping track of which autopsies are on which tape should not be difficult as all the autopsies are stored in ascending order by post-mortem number, which will provide the person doing updating and error correction immediate access to the correct tape.

5.1 Data Base Structure

The design of the autopsy data base is very similar to the full output format described in Chapter IV; however, due to the several differences and the importance of having every item correctly specified, the structure will be described here in detail.

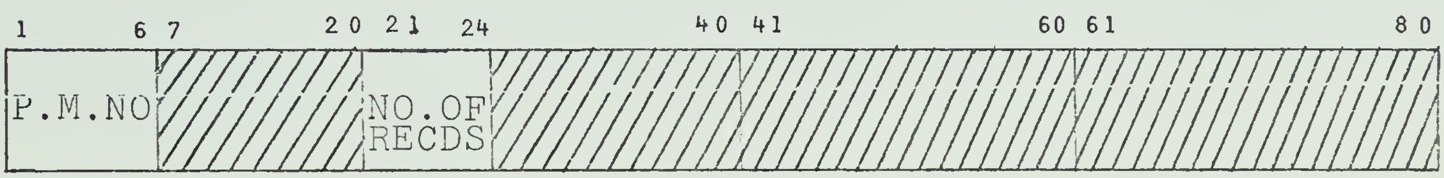
The autopsy reports are stored consecutively in order of post-mortem number, i.e. by year and sequence number within that year. Every record on the autopsy tape is a card image, i.e. 80 characters long. In most instances, each record may be sub-divided into four 20-character fields as follows.

1 — 20	21 — 40	41 — 60	61 — 80
FIELD1	FIELD2	FIELD3	FIELD4

Each autopsy report is composed of five sections: general information, pathological diagnosis, measures, drugs and other therapy, and medical history. In the following description usage of the terms 'record', 'line', or 'card' are all synonymous. In order to avoid repetition, the reader will often be referred back to Chapter III, Question Formulation, for the details of any formats previously described.

5.1.1 General Information Division

The first line of each autopsy report contains the post-mortem number in FIELD1 and the number of records to follow within the autopsy report in FIELD2. The post-mortem number differs from the format previously described in that the hyphen is omitted and thus only requires 6 characters; these must be left-justified. The number-of-records entry must be right-justified in the first 4 positions of FIELD2, for example 30 would be specified as b̸b̸30, where b̸ represents a blank. This first line is used for ordering and locating purposes in the updating and error-correcting procedures.



The second line contains the case name, post-mortem number, University of Alberta Hospital patient identification number, and the version date in FIELD1, FIELD2, FIELD3, and FIELD4, respectively. The name and post-mortem number must be left-justified and be of the same format as described in Chapter III. The patient identification number must be 6 digits long (inserting leading zeros if necessary) and must be left-justified in its designated field. The version date must conform to the date specification standards adopted for all portions of this System, and shows the last time any changes were made to this particular autopsy report.

1 16		21 27	28 40	41 46	47 60	61 70	71 80
NAME		P.M.NO.		UAH NO.		VERSION DATE	

The third line includes the age, sex, and race of the deceased person in FIELD1, FIELD2, and FIELD3 respectively, and this is followed on the next line by the birth date, date of admission to the hospital, death date, and autopsy date in that order. The contents of these fields have either been described previously or conform to adopted standards.

1 3	4 20	21	22 40	41	42 60	61 80
AGE		SEX		RACE		

1 10	11 20	21 30	31 40	41 50	51 60	61 70	71 80
BIRTH DATE		ADMIT DATE		DEATH DATE		AUTOPSY DATE	

Line 5 of every autopsy report contains the elapsed time from death to the time of the autopsy as well as the completeness of the autopsy. The format of this record varies slightly from the normal format - FIELD1 contains an abbreviation of the form 10H (10 hours), 2D (2 days), or 1W (1 week) left-justified to indicate the time elapsed; FIELD2, FIELD3, and FIELD4 are combined to form one field containing any statement regarding the completeness of the

autopsy, for example COMPLETE, or PARTIAL - CHEST ONLY.

1 20	21 80
ELAPSED TIME	COMPLETENESS OF AUTOPSY

The next line shows the surnames of the doctors associated with this case, FIELD1 being for the resident pathologist, FIELD2 the staff pathologist, and FIELD3 the service doctor. The surnames must be left-justified and limited to a maximum of 20 characters each.

1 20	21 40	41 60	61 80
RESIDENT PATHOLOGIST	STAFF PATHOLOGIST	SERVICE DOCTOR	

The last portion of the general information division contains the clinical diagnosis. The clinical diagnosis may extend for as many lines as required; however, each line may only contain one word in each of the four fields. Each of the four words per card must be left-justified within its respective field and truncated to a maximum of 20 characters.

1 20	21 40	41 60	61 80
CLINICAL DIAGNOSIS TERM	CLINICAL DIAGNOSIS TERM	CLINICAL DIAGNOSIS TERM	CLINICAL DIAGNOSIS TERM

The end of the general information division is signified by a '9' in column 1 of the line following the clinical diagnosis, the line being left blank otherwise.

5.1.2 Pathological Diagnosis Division

The pathological diagnosis division will normally consist of a large number of statements, all of the same format, describing the findings of the post-mortem examination. FIELD1 contains the name of a body part or an accepted exception such as WOUND or OPERATION. This field must always be filled, its contents left-justified and truncated to 20 characters. The adjectives LEFT and RIGHT also appear in this field, in the format described in Chapter III. The remaining three fields of each record contain the actual findings and are formatted one word per field, with that word being left-justified and truncated in the usual manner. If a body part requires more than three words to describe the findings, an unlimited number of extensions are allowed by simply repeating the body part name in FIELD1 and continuing the description in the following fields. In the full-output option described in the previous chapter, the body part name appears only once; however, when constructing the data base it must be repeated for each record.

1 20	21 40	41 60	61 80
BODY PART	DESCRIPTOR	DESCRIPTOR	DESCRIPTOR

Again, a '9' in column 1 of a blank line marks the end of the division. If there were an instance where there is no information to record in the pathological diagnosis division, this delimeter must still

occur to signify the fact; absence will result in the measures division being interpreted as part of the pathological diagnosis division and everything thereafter similarly displaced.

5.1.3 Measures Division

The measures division, which contains weights and volumes of body organs and fluids, makes use of all four fields of each record by storing two entries per line. FIELD1 contains the name of an object measured and FIELD2 contains the measure; FIELD3 contains the name of a second object that was measured and FIELD4 contains its measure. The name field may contain a maximum of 20 characters, which must be left-justified and conform to the specifications outlined in Chapter III for searching this particular entity. The measure field contains exactly 10 characters in the manner described in the same chapter.

1 – 20	21 – 30	31 – 40	41 – 60	61 – 70	71 – 80
OBJECT MEASURED	MEASURE		OBJECT MEASURED	MEASURE	

Once again a '9' in column 1 of the last line of this division informs the system that the end has been reached and to anticipate the drug and other therapy division.

5.1.4 Drugs and Other Therapy Division

Each instance of therapy administered to a patient requires three lines within the data base for storing all the required information. The first line contains the therapy name in positions 1-20, the dosage in positions 21-30, the rate of administration in positions 33-40, the route of administration in positions 41-44, the starting date in positions 45-54, and finally the ending date in positions 57-66. The correct method of specifying information within each of these fields, as well as the remaining two lines for each instance of therapy, was discussed in Chapter III.

1 20	21 30	31	33 40	41 44	45 54	55 56	57 66	80
THERAPY NAME	DOSAGE		RATE	ROUTE	START DATE		END DATE	

The second line contains the reason that the previously defined therapy was administered. The reason for administration must be stated in a maximum of four words, each restricted to 16 characters, and so confined to a single line. The therapy name is not required in this record since the proper association is automatically made with the therapy named in the preceding line.

1 16	17 20	21 36	37 40	41 56	57 60	61 76	77 80
REASON TERM		REASON TERM		REASON TERM		REASON TERM	

The third line of each triple contains any sequel that may be attributed to the administration of the therapy in question; it is constructed in the same manner as the previous line.

1 — 16	17 — 20	21 — 36	37 — 40	41 — 56	57 — 60	61 — 76	77 — 80
SEQUEL TERM		SEQUEL TERM		SEQUEL TERM		SEQUEL TERM	

The three lines for an instance of therapy are consecutive in the order indicated, and are followed by further sets of three as appropriate.

As previously, the standard delimiter of a '9' in column 1 of a blank line identifies the end of the division and must be present.

5.1.5 Medical History Division

The medical history division contains one type of record format to store any significant medical occurrence during a patient's life. Each line is constructed as follows: date in positions 1-10; sequence number in positions 13-15; and positions 21-80 contain three 20-character fields in which the incidents of the given date are recorded. The date representation is in the standard format adopted throughout the System. The sequence number, which commences at 1 and is incremented by 1 for each line with the same date, should be left-justified in its desig-

nated field. The three fields devoted to describing the occurrences permit one word per field, in the format described in Chapter III; if the allotted space is not sufficient, one may continue on as many lines as required providing each line cites the date and a sequence number incremented by 1 from the previous line. Besides controlling multiple cards for one occurrence, the sequence numbers are useful where two or more significant medical happenings occur on the same day and the order of occurrence is important. In the previously described full-output option, a date appears only once and thereafter only the sequence numbers and events are output; however, in the data base every record must contain a date and sequence number.

1 10		13 15	16 20	21 40	41 60	61 80
DATE		SEQ NO.		EVENT	EVENT	EVENT

Again the division concludes with a blank line containing a '9' in column 1.

5.1.6 Conclusion of the Data Base

Immediately following the conclusion of one autopsy report the next one is begun, i.e. the medical history division of one report is followed by the general information division of the next report, with the '9' at the end of the medical history division acting as the delimiter between the two reports. This process continues until the

contents of the last autopsy report is recorded. At this point two additional lines, signifying the end of the data base, are inserted. The first contains '99' in positions 1-2. The second contains, in positions 1-10, the date, in the previously defined format, of the most recent update of any portion of the autopsy data base. It is this date that appears on the same page as the question listing in either of the search output formats. It is imperative that this date be updated for even the most minor change to the data base.

The following illustration shows the general design of a typical autopsy report.

```
GENERAL INFORMATION DIVISION
9
PATHOLOGICAL DIAGNOSIS DIVISION
9
MEASURES DIVISION
9
DRUGS AND OTHER THERAPY DIVISION
9
MEDICAL HISTORY DIVISION
9
```

If this were the concluding autopsy report in the data base, the following two lines would be appended.

```
99
1971/06/16     (for example)
```

5.2 Updating the Autopsy Data Base

Updating the autopsy data base would probably be done once a month, in compliance with the general procedures

within the Department of Pathology, whereby each pathologist is assigned the task of performing autopsies for a one-month period on a rotating basis. At the conclusion of his tour of duty, the pathologist usually devotes the following month to formally documenting the post-mortem report, after which it is ready for coding and insertion into the Autopsy Storage and Retrieval System.

From the user's view of the complete updating procedure, the operation seems to be fairly simple, only requiring three stages; however, if one examines that part of the procedure not readily visible to the user, one realizes the complexities involved. The updating procedure was designed so as to keep the amount of work and computer knowledge required by the person assigned to this duty at a minimum.

5.2.1 Updating: A User's View

Prior to outlining the complete updating procedure, the instructions required by the user to update the tape file will be described. The three stages involved in this operation are - (1) creating a file with the new autopsy reports that are to be added; (2) correcting any errors that may exist in this file; (3) doing the actual updating onto the autopsy data tape.

The recommended method for entering the new data into a file is via cards, since there would probably be a sub-

stantial amount of data and the cost of connect time while sitting at the terminal keying in the information could be considerable. In the use of cards, errors are far easier to correct prior to storing the data in a file. The autopsies should be punched on cards, in the format described at the beginning of this chapter, omitting that portion of the first line which contains the number of lines to follow in the report. At the conclusion of all the new autopsy reports, there must appear the same two lines as would be found at the end of the data base, i.e. '99' in the second to last record and the current date in the last one. Assuming that all the information is correctly keypunched onto cards, the user would submit the following batch job.

```
$SIGNON  etc.
Password
$EMPTY TEMP
$GET TEMP
$NUMBER
          }    the new data in ascending
               post-mortem number order
$UNNUMBER
$RUN DATACHK 1=TEMP 2=-RECNUMB 3=-A
$LIST TEMP
$SIGNOFF
```

The new data is placed in TEMP, which is a file in permanent existence on disk housing the updating information. This file contains the autopsy reports from the previous updating and must be emptied prior to being refilled.

After the new autopsy data is entered into TEMP, the program DATACHK is executed. This program inserts the

record count in the first line of each autopsy report, checks that the final two lines are present, and checks that the correct number of division delimiters are present, i.e. the lines with the '9' in character position 1. If the new data satisfies the checks performed by DATACHK, the message NO DETECTED ERROR IN INPUT DATA is output on the page preceding the listing of TEMP. If the final two lines expected in TEMP are absent, the diagnostic message is ERROR IN INPUT DATA - PROBABLY OMISSION OF 99 AND DATE LINES. If a division delimiter should be omitted, one of the following two diagnostics will appear: ERROR IN INPUT DATA - PROBABLY OMISSION OF A DIVISION DELIMITER or FCVTH - INVALID CHARACTER IN NUMERIC FIELD. Depending upon where the delimiter was omitted, the error may be discovered by DATACHK, resulting in the former diagnostic, or the operating system, resulting in the latter. In most instances it will be the operating system discovering the error, since DATACHK only determines if the correct number of delimiters are present after having read the complete file. If any of the above mentioned errors should be found in TEMP, the record count will not be inserted in the first line; however, execution of the final step continues. In this case it will be necessary to reexecute DATACHK upon the completion of editing.

The next stage is that of editing TEMP. Upon examination of the contents of TEMP, any errors noticed, whether due to omissions, keypunching, or other reasons, should be

corrected. There are two equally acceptable methods by which one may correct the file - making manual changes to the card deck and rerunning the batch job, or using the MTS file editing facilities. The only drawback to the first is waiting for the execution of the batch job and obtaining the output listing; however, there is the advantage of not having to become involved with the file editor.

If the latter choice is taken, and assuming the user is signed on, the command

$RUN *EDIT

will enter him into the editing mode. The MTS system will then prompt the user for the name of the file to be edited. The user should respond with the name TEMP. The editing commands most likely to be used in the editing step are the following.

COLUMN
OVERLAY
BLANK
REPLACE
INSERT

A complete description of these commands, as well as other available ones, can be found in the MTS File Editor manual published by the University of Alberta Computing Center. Upon completion of all the desired editing operations, the STOP command will return the user to the MTS command mode and allow him to continue with the final stage. The echo

facility of the file editor is very handy in that it allows the user to see whether he has made the corrections properly.

The third stage in the updating procedure is constructed in a manner similar to the searching procedure, for the sake of simplicity. All the MTS commands required for the actual updating reside in the file UPDTPROC, and the user need only direct the operating system to this file for its instructions. The best method of initiating this step of the procedure is as a *BATCH job, i.e. a batch job submitted from the terminal. Assuming the user is signed on at the terminal, the following series of commands will complete the updating.

```
$RUN *BATCH
$SIGNON    etc.
Password
$SOURCE UPDTPROC
$ENDFILE
```

At this point the user may sign off from the system or proceed to do anything else he desires. Until such time as the user is signed off, the *BATCH job will not begin executing, unless of course the sign-on identification numbers differ in both cases.

Since the updating procedure makes use of the file QUESTIONS and empties its contents prior to using it, no attempt to enter search questions should be made during the updating process, nor should questions be appended to the end of the file.

The update process produces a printed record of all autopsy reports affected, in the same style as the full-output option from a submitted question. The System generates questions from the file TEMP and then searches TEMP using these questions. The result of the search is one autopsy report for each question. The output is used for keeping the records of the Department of Pathology up-to-date and for mailing results from an autopsy to the service doctor involved.

5.2.2 Updating: A System's View

The UPDTPROC file contains all the instructions for the actual updating, the question creation, and the preparation of output. The contents of this file are listed below and an explanation of the process follows thereafter.

```
 1. $RUN *MOUNT PAR=0034 9TP *TAPE* VOL=T00034 SIZE=7280
          LRECL=80 FMT=FB RING=IN
 2. $RUN *MOUNT PAR=0151 9TP *SCRATCH* VOL=T00151 SIZE=7280
          LRECL=80 FMT=FB RING=IN
 3. $RUN *TAPECOPY 0=*TAPE* 1=*SCRATCH*
 4. $RUN *DISMOUNT PAR=*SCRATCH*
 5. $RUN *MOUNT PAR=0141 9TP *BAKUP* VOL=T00141 SIZE=7280
          LRECL=80 FMT=FB RING=IN
 6. $RUN UPDATE 1=TEMP 2=*TAPE* 3=*BAKUP*
 7. $RUN *TAPECOPY 0=*BAKUP* 1=*TAPE*
 8. $EMPTY QUESTIONS
 9. $RUN MAILREC 5=TEMP 7=QUESTIONS
10. $CREATE -SORTIN SIZE=100P
11. $RUN INQUIRY 1=QUESTIONS 5=TEMP 6=-SORTIN
12. $CREATE -SORTOUT SIZE=100P
13. $RUN *SORT
14. SORT=CH;A;121;9;CH;A;130;4
15. INPUT=-SORTIN;F;133;133
16. OUTPUT=-SORTOUT;F;133;133 MNR=50000
17. $RUN *EDIT SPRINT=*DUMMY* GUSER=*SOURCE*
18. -SORTOUT
```

```
19. REGION /A 1 *L
20. COL 121 133
21. B /A '          '
22. STOP
23. $COPY -SORTOUT *SINK*
24. $SIG
```

The first 7 lines of this file constitute the actual updating. There are three tapes involved in this process, only two of which are mounted on drives at any one specific time. *TAPE* contains the autopsy data base, *BAKUP* is an exact duplicate of *TAPE* kept as a backup tape, and *SCRATCH* is a scratch tape. In order to avoid having the user issue tape mounting commands, a small amount of juggling had to be done so that upon completion of the updating process, all the tapes retain their original rack number and purpose. All tapes are identified to the computer operator by their storage rack number and the intention of the Autopsy System is to have rack number 34 always containing *TAPE*, i.e. the autopsy data base, and similarly for the other tapes being used by the System. This method, although slightly confusing, leaves the name of a tape and its associated rack number unaltered, thus permitting the tape mount commands to remain unchanged. The method is to copy *TAPE* to *SCRATCH* and dismount the latter, which temporarily becomes the backup tape. *BAKUP* is then mounted and the program UPDATE is executed. This program reads *TAPE* and the file TEMP and merges them in acending post-mortem number order onto *BAKUP*. Finally *BAKUP* is copied

to *TAPE* and each tape returns to serving its original intention as implied by the name assigned to it. The following flow diagram should clarify any confusion within the reader's mind. The numbers in brackets refer to the order in which the actions take place, and the numbers within the boxes refer to the logical device numbers assigned to each unit.

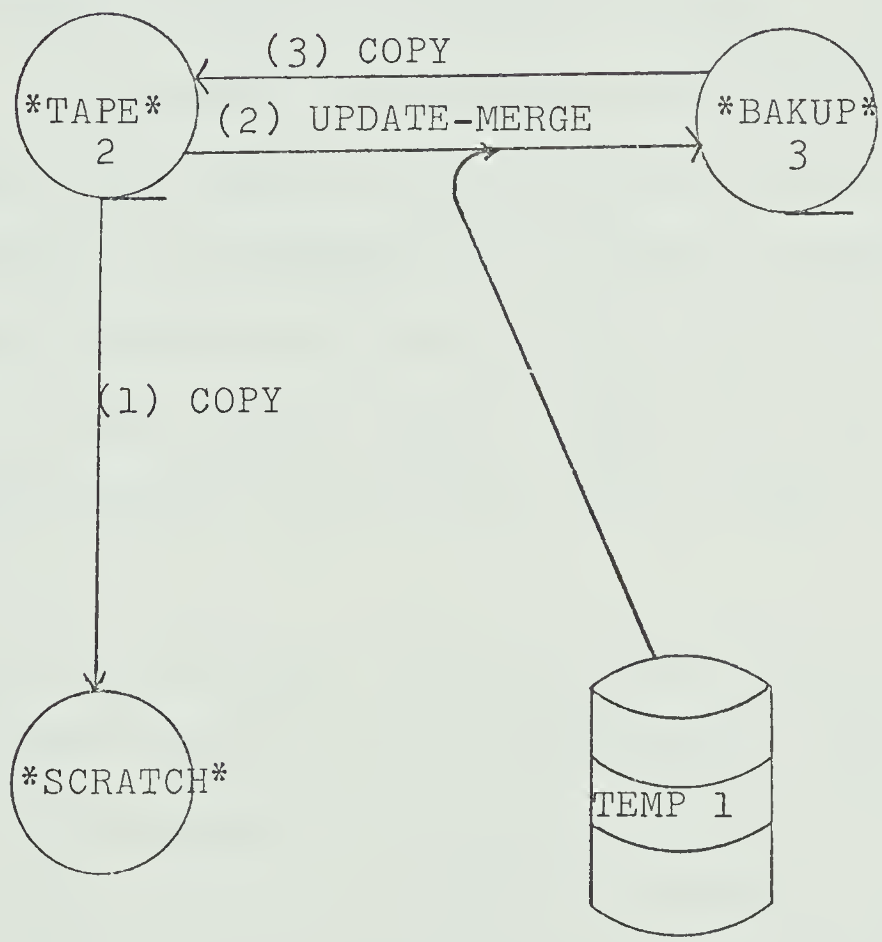


Fig. 5.1: Updating Procedure

Line 8 empties the contents of the file QUESTIONS in preparation for the output from the program MAILREC, which is called for execution in the next line. MAILREC scans the file containing the new autopsy reports, i.e. TEMP,

extracts the name and post-mortem number of each case, and formulates a question in the file QUESTIONS for each of these new reports. The form of the composed question is as follows.

```
QUE      UPDATE/ERROR CORRECTION OUTPUT          1
AND N    name
AND PE   post-mortem number
```

The remainder of UPDTPROC is the same as the previously described SEARPROC, i.e. searching procedure, with one exception. In SEARPROC the contents of QUESTIONS are compared against the complete autopsy data base residing on *TAPE*; however, in UPDTPROC the contents of QUESTIONS are compared only against the new autopsy reports in TEMP. There will be one autopsy report output for each question in QUESTIONS and one question for each autopsy report in TEMP.

If anything untoward should occur during the updating procedure which results in an irrecoverable error on *TAPE* and loss of the complete data base, a copy of the data base will be found on either *SCRATCH* or *BAKUP*, depending at which point in the updating procedure this error occurred. If the backup resides on *BAKUP* then one need only copy *BAKUP* to *TAPE* and restart the updating process from the third stage. If the backup version resides on *SCRATCH*, then this must be copied to *TAPE* and *BAKUP*, only after which can updating continue from the third stage.

5.3 Data Base Error Correction

Regardless of how closely one checks the contents of TEMP in the updating process, mistakes in the data base will inevitably be discovered at a later time. In anticipation of this occurring, an error correction procedure has been devised to allow the user to set right any misrepresented information on the autopsy data base tape. Once again, in an attempt to simplify operations for the user, most of the error correcting procedure is not directly visible. It would be costly to correct every error immediately on discovery, and so it would be advisable to note the post-mortem number and the complete content of the line(s) in error, then make corrections monthly or so.

5.3.1 Error Correction: A User's View

All the work done by the user in performing corrections will be done at the terminal. Assuming the user is signed on and has available a list of where corrections are to be made, the first step requires creation of the file AUTNOS which will hold the post-mortem numbers of the autopsy reports to be corrected. Each line of this file should contain only one post-mortem number in the format that it appears in the first record of every autopsy report, i.e. without the hyphen. The contents of AUTNOS must be ordered in ascending post-mortem number order. Upon concluding the entry of post-mortem numbers into AUTNOS, this complete file

should be listed so that the user may verify the correctness of each entry and make certain that none have been duplicated or omitted. The following are the MTS commands required for performing the above described operations.

```
$CREATE AUTNOS
$NUMBER
                        post-mortem numbers
$UNNUMBER
$LIST AUTNOS
```

If any errors are discovered in the listing of AUTNOS, the next step would be to edit the file, making the required changes. Once the contents of AUTNOS are considered to be correct, the user directs MTS to the file ERRPROC for commands to continue the error correcting process. This is done as follows.

```
$SOURCE ERRPROC
```

After ERRPROC performs certain operations, to be described in the next section, it returns control to the user at the terminal who will now perform the actual corrections. At this point all the autopsy reports that require corrections will reside in the file CORR in ascending post-mortem number order, which would also be the best order for making corrections. In order to make changes to CORR, the MTS file editor must be invoked as follows.

```
$RUN *EDIT
CORR
REGION /A 1 *L
                        correcting commands
STOP
```

In order to determine the location of a particular line within CORR that is to be corrected, use of the file editing command SCAN is recommended. For example, if the beginning of a particular record to be corrected contains the term LIVER, one could easily determine exactly where it is located by using the SCAN command as follows.

```
SCAN@A /A 'LIVER'
```

This informs the editor to scan the region A, previously defined to be the complete file, for all records beginning with the string 'LIVER'. All lines meeting this criterion will be printed at the terminal with their respective line numbers within CORR. From this list the exact record and line number can be determined the user may proceed to use other editing commands to correct the line. This process is continued until all the required corrections are made to CORR. The user is reminded that the version date on the second line of every autopsy report must be appropriately updated every time any change is made to any part of the report. If any of the corrections involve the

addition or deletion of lines, the user must make the appropriate adjustment to the record count entry on the first line of the autopsy report.

Once editing is complete the user must add two more lines to the end of CORR, the same as those appearing at the end of the data base, i.e. '99' and the current date. The following instructions will allow the addition of the two records.

```
$GET CORR
$NUMBER LAST+1
99
1971/06/16      (for example)
$UNNUMBER
```

At this point the user has completed all the work he must do in the error correcting procedure other than issuing the following command to return control to ERRPROC so that it may continue where it left off.

```
$SOURCE PREVIOUS
```

The user must wait at the terminal to obtain the receipt number from the *BATCH job which is to be submitted by ERRPROC, which will also automatically sign him off.

5.3.2 Error Correction: A System's View

Although the user must do the actual file editing to correct the errors, the majority of the work involved in the error correcting procedure is not readily visible to

the user but contained in ERRPROC. The contents of this file is listed below and a discussion follows that.

```
 1. $RUN *MOUNT PAR=0034 9TP *TAPE* VOL=T00034 SIZE=7280
          LRECL=80 FMT=FB RING=OUT
 2. $RUN *MOUNT PAR=0151 9TP *SCRATCH* VOL=T00151 SIZE=7280
          LRECL=80 FMT=FB RING=IN
 3. $CREATE CORR SIZE=30P
 4. $RUN ERROR 1=*TAPE* 2=AUTNOS 3=CORR 4=*SCRATCH*
 5. $RUN *DISMOUNT PAR=*TAPE*
 6. $RUN *DISMOUNT PAR=*SCRATCH*
 7. $SOURCE *MSOURCE*
 8. *RUN *BATCH
 9. $SIG  etc.
10. Password
11. $RUN *MOUNT PAR=0034 9TP *TAPE* VOL=T00034 SIZE=7280
          LRECL=80 FMT=FB RING=IN
12. $RUN *MOUNT PAR=0151 9TP *SCRATCH* VOL=T00151 SIZE=7280
          LRECL=80 FMT=FB RING=OUT
13. $RUN UPDATE 1=CORR 2=*SCRATCH* 3=*TAPE*
14. $RUN *DISMOUNT PAR=*SCRATCH*
15. $DESTROY AUTNOS
16. $EMPTY QUESTIONS
17. $RUN MAILREC 5=CORR 7=QUESTIONS
18. $CREATE -SORTIN SIZE=100P
19. $RUN INQUIRY 1=QUESTIONS 5=CORR 6=-SORTIN
20. $CREATE -SORTOUT SIZE=100P
21. $RUN *SORT
22. SORT=CH;A;121;9;CH;A;130;4
23. INPUT =-SORTIN;F;133;133
24. OUTPUT=-SORTOUT;F;133;133 MNR=50000
25. $RUN *EDIT SPRINT=*DUMMY* GUSER=*SOURCE*
26. -SORTOUT
27. REGION /A 1 *L
28. COL 121 133
29. B /A '             '
30. STOP
31. $COPY -SORTOUT *SINK*
32. $DESTROY CORR
33. $SIG
34. $ENDFILE
35. $SIG
```

The first six lines of this file contain procedures involved with running of the program ERROR. This program compares the post-mortem numbers of the autopsy reports to

be corrected against the post-mortem numbers on the complete autopsy data base tape, i.e. AUTNOS is compared against *TAPE*. If a post-mortem number on *TAPE* is identical to a post-mortem number in AUTNOS, the complete autopsy report is copied to the file CORR for correction; if the post-mortem numbers do not match, the complete autopsy report is written onto *SCRATCH*. When ERROR completes its task the original date base is split into two files - *SCRATCH* containing the original data base minus those autopsy reports requiring correction, and CORR, which contains the latter. The operations performed by ERROR are illustrated in the following flow diagram. The numbers in brackets denote the order in which events occur.

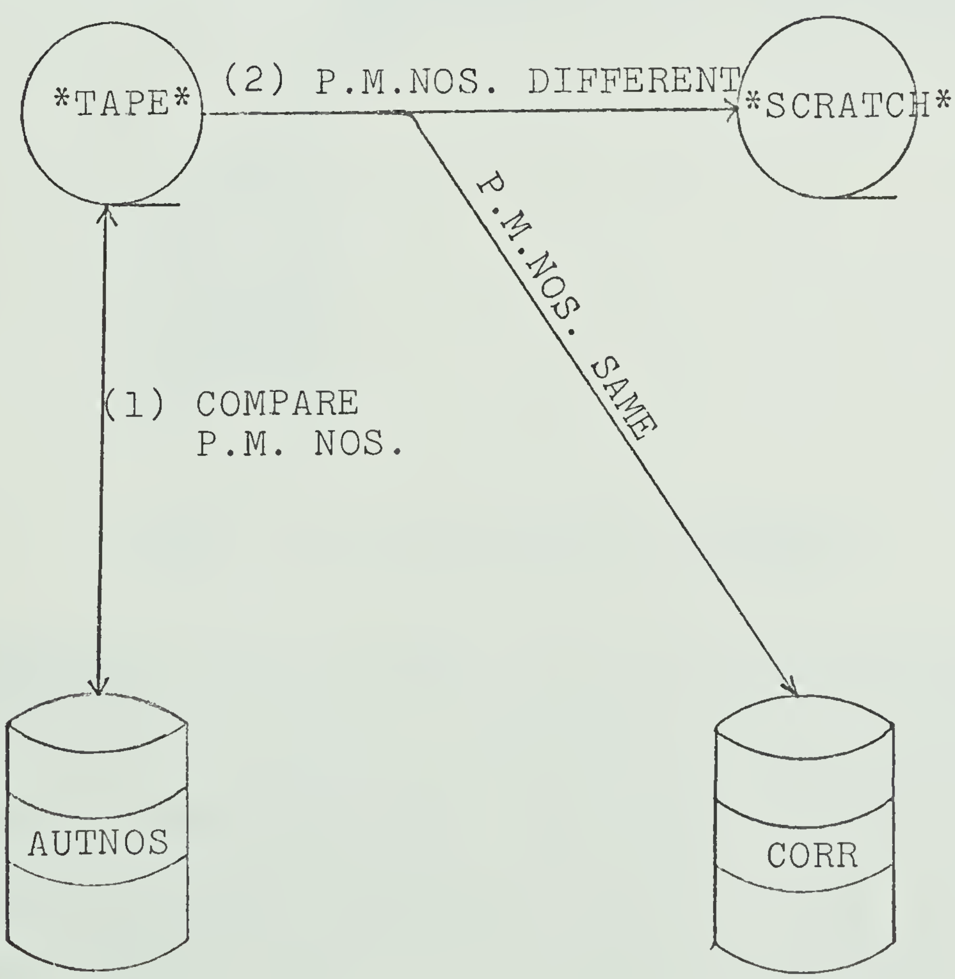


Fig. 5.2: Processing in ERROR

The command in line 7 transfers control back to the user at the terminal, at which point he commences editing to set right the errors in CORR. Upon completion of the editing, the user transfers control back again to ERRPROC, which now submits the remainder of the data base correction procedure to MTS as a *BATCH job.

The first step of this *BATCH job executes the program UPDATE, which merges the now corrected file CORR and *SCRATCH* to form the new error-free data base on the tape *TAPE*.

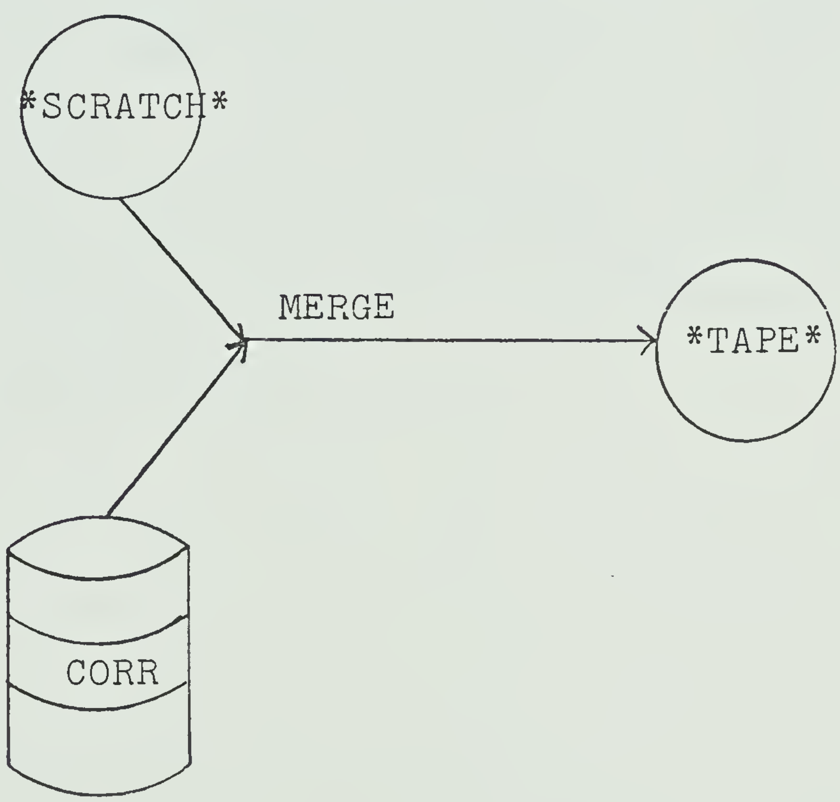


Fig. 5.3: Processing in UPDATE

The remainder of the *BATCH job is the same as the procedure followed in updating; MAILREC creates questions from the data in CORR and places them into the file QUESTIONS, the search program INQUIRY is executed using CORR as the data base against which this search is performed, then

final output from the data base error-correcting procedure is prepared as the full-output option for each of the created questions.

In the complete error-correcting process nothing is ever done to update the backup tape *BAKUP*. In the event that something untoward should occur during error correction which results in the loss of the complete data base on *TAPE*, the backup tape *BAKUP* can be copied to *TAPE* and the situation would be as if no error correction had been done. The only time *BAKUP* is updated is during the updating procedures. If this type of error should occur the best thing to do after having recreated the data base is to destroy AUTNOS and CORR and to recommence the complete error correction routine.

CHAPTER VI

THE DATA BASE SEARCHING ALGORITHM

The method of formulating and submitting a question to the Autopsy System was discussed in Chapter III, and the autopsy data base described subsequently. This chapter will concentrate on a description of the algorithm, as developed by Professor H.S. Heaps and Analyst L.H. Thiel of the Department of Computing Science at the University of Alberta and extended by the research described in this thesis, implemented to perform the searching of the autopsies within the data base.

The searching program is written entirely in Fortran and the object module of this program, as compiled by the Fortran H compiler for the IBM 360/67, resides in the permanent disk file INQUIRY. The program is composed of five routines - MAIN, SEARCH, LOCATE, COMPAR, and PAGE. Each of these routines will be outlined in the sections of this chapter. The searching algorithm only requires that the data base be read once for each batch of questions, regardless of the number of questions within the batch. (The batch has, as stated, been limited to ten questions, because of storage constraints).

6.1 The MAIN Line

The MAIN routine analyzes the questions submitted by the user and prepares tables to allow for the comparison with the data base. The result is the creation of three tables - the TERM table, POSITION table, and QUESTION-PARAMETER table.

The following figure illustrates the construction of the TERM table.

	TERM										POINTER	CODE
	T(1)	T(2)	T(3)	T(4)	T(5)	T(6)	T(7)	T(8)	T(9)	T(10)		
1	ANAP	LAST	IC								3	7
2	CARC	INOM	A								21	7
3												
	.	.	.								.	.
	.	.	.								.	.
	.	.	.								.	.
	.	.	.								.	.
99												
100												

Fig. 6.1: The TERM Table

Each unique term has an entry in this alphabetically ordered table. The 'term' portion of the TERM table is divided into ten divisions of four characters each to allow a maximum of 40 characters, which is equivalent to the maximum amount of space allotted to term specification on the TERM card. In most instances additional predefined characters are added to the given term so as to aid in identifying which portion of the autopsy report is to be searched for the term. The pointer field points to an entry in the POSITION table, which provides further information about the term. The code field contains a number from 1-7 to identify the type of comparison to be made with the data base, i.e. inequality

search or not, and in the case of an inequality search the location of the term in the term card. The interpretation of the codes is as follows.

1 : ≤TERM1
2 : =TERM1
3 : ≥TERM1
4 : =TERM1 and ≤TERM2
5 : =TERM1 and =TERM2
6 : =TERM1 and ≥TERM2
7 : =TERM1 and/or =TERM2

The maximum length of the TERM table is 100 entries, meaning that no batch of questions may contain more than 100 term cards in total.

The design of the POSITION table is shown in the following illustration.

	QUESTION NO.	PARAMETER NO.	BACK POINTER
1			
2			
3	2	3	-1
	.	.	.
	.	.	.
12	2	2	-1
	.	.	.
	.	.	.
	.	.	.
21	5	1	12
	.	.	.
	.	.	.
	.	.	.
	.	.	.
99			
100			

Fig. 6.2: The POSITION Table

Unlike the TERM table, which contains an entry for each unique term, the POSITION table contains an entry for every request of a term; if "cancer" appears three times in the same context within a batch of questions, the TERM table will have one entry for cancer, the POSITION table will have three. The first column of the POSITION table gives the question number wherein the term appears, the second provides the parameter number of the term within the question. Each occurrence of AND or NOT in the logic field increments the number of parameters within the question by one. Since the pointer column of the TERM table relates to only one entry within the POSITION table, but a term may have multiple entries therein, there must exist a scheme for linking identical terms within this table. The 'back pointer' column provides this by pointing to a further entry for the same term within the table. An entry of '-1' in this column signifies that the end of the chain has been reached.

The QUESTION-PARAMETER table is only half filled in this initial routine with the remainder being completed in further routines. Its structure is as follows.

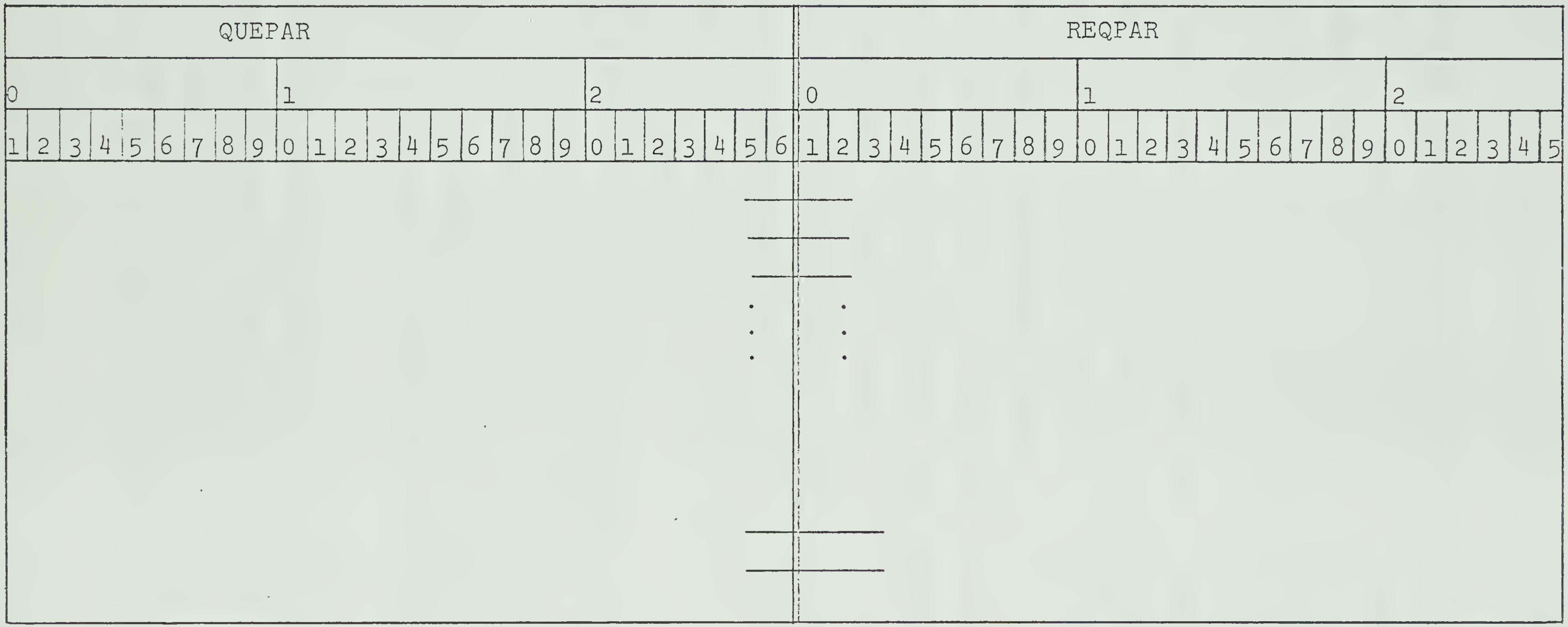


Fig. 6.3: The QUESTION-PARAMETER Table

The QUESTION-PARAMETER table is composed of two Boolean matrices (i.e. each entry is either 0 or 1) named REQPAR and QUEPAR, the former having 25 columns and the latter 26. Each matrix has one row devoted to each question within a batch. The first routine in the searching algorithm completely fills REQPAR as well as the 26th column of QUEPAR. Each question may have a maximum of 25 parameters, i.e. AND plus NOT cards, and space is provided for one entry for each of these possible parameters within each line of REQPAR. Each AND parameter has a '1' placed in its corresponding parameter position within REQPAR; a NOT parameter has a '0' similarly placed. The 26th column of QUEPAR contains the output format information extracted from the 80th character position of the question card. A '1' in column 26 signifies that the full-output option is being requested, a '0' signifies abbreviated output.

In addition to creating and entering information into the previously defined tables, the MAIN line checks for format errors within a question, e.g. a field mispositioned or a required blank omitted. If an error should be discovered, a diagnostic message is generated and the question invalidated.

A further task performed by this routine is the assignment of sequence numbers to each line in the question batch so that ordering is retained in output. The increment of 1,000,000 numbers between the beginning of successive

questions is more than sufficient to allow space for retrieved autopsy reports to be inserted in sequence with the questions they are associated with.

6.2 Subroutine SEARCH

The main functions of the SEARCH routine are reading the data base and effecting output. The records on the autopsy data base tape are read one at a time. After each record is read, control is passed to one of two routines to perform the actual comparisons with the entries in the TERM table. These routines, which will be described in the following two sections of this chapter, also fill in the QUEPAR matrix. At the completion of reading an autopsy report, SEARCH compares the QUEPAR and REQPAR matrices for each individual question within a batch, omitting the 26th column of QUEPAR. For each case where these are exactly the same, the autopsy report is copied into a file as part of the answer for that question. Prior to commencing the reading of each autopsy report, the QUEPAR matrix, excluding the 26th column, is reset to zero so that the same process may be repeated.

In terms of the number of Fortran statements, SEARCH is the longest routine within the System; however, this is attributed to the fact that the majority of the statements serve to read the data base and for the output of autopsy reports that meet question requirements.

6.3 Subroutine LOCATE

Subroutine LOCATE, which is one of two routines that perform comparisons, performs two tasks within the overall System. The MAIN line, which creates the TERM table from the terms within submitted questions, calls this routine to assure that each term has only one entry within the table and to order the table alphabetically. The second purpose that this routine serves is to perform equality search comparisons, as for example searching the pathological diagnosis division. In this capacity it is called by the subroutine SEARCH, and immediately transfers control back to SEARCH upon completion of its task. Having performed the comparison between an entry in the TERM table and a portion of the data base, LOCATE fills the QUEPAR matrix of the QUESTION-PARAMETER table. If a match is encountered in the comparison, the POSITION table is referenced to determine all occurrences of this term within the batch of questions, and a '1' entered in the appropriate positions of QUEPAR.

6.4 Subroutine COMPAR

This routine, similar to LOCATE, is also called by SEARCH; its purpose is to perform inequality search comparisons. There are two types of inequality searches that this routine handles - those with codes 1-3 and those with codes 4-6. In the first instance, the inequality search comparison is applied against the item in the TERM1 field of the

term card. In the second case, an equality against the contents of TERM1 must be established initially, followed by an inequality search on the contents of TERM2. If an item within the data base satisfies a required inequality, QUEPAR is appropriately updated in the manner previously described.

6.5 Subroutine PAGE

The final subroutine within the searching algorithm is responsible for the page header appearing at the top of each output page, other than those containing the listing of a question. Since the System is unaware of the length of an autopsy report when it is being output, a line count is kept within the subroutine SEARCH. As the end of a page may occur after any line of output, approximately 25 Fortran statements would have to be inserted after each record is written so as to make certain that the page header appears at the top of each page. Rather than inserting this large amount of code in SEARCH, the statements required to display the page header are located in subroutine PAGE, and a call to this subroutine inserted after the writing of each record. SEARCH will only transfer control to this subroutine if the line count has reached a prescribed value.

CHAPTER VII

SUMMARY

The objective of the research presented in this thesis was to provide the Department of Pathology at the University of Alberta Hospital with the ability to use a computer for the storage and retrieval of post-mortem examination reports. The main argument for a system of this nature was to facilitate research within the Department. Previously, a manual search through the massive files of the Department was necessary and much valuable time was expended. With the implementation of the University of Alberta Hospital Autopsy Storage and Retrieval System, a user need only specify his desires in a formatted question and the computer takes over to search all the autopsy reports in the computer tape file.

The searching algorithm employed allows the free usage of English words in the searches, as opposed to coding the questions and the data base. There does exist a pathology coding scheme, called Systematized Nomenclature of Pathology (SNOP); however, after examination of the code it was found not to be very suitable for this application. The SNOP coding scheme contains a large vocabulary and the coding from English to SNOP and vice versa would require too much time and money to be warranted.

A possible extension of the research herein described would be the development of a short, simple, and fast algo-

rithm for coding words in the English language [13,17]. This scheme should be able to take any combination of letters used to form English words and with the aid of a computer code the letters and then have the ability to decode the data without the loss of any information. Applying such a coding algorithm to the Autopsy System would not alter any appearances from the user's standpoint; however, there would be gross internal changes. Autopsy reports and questions would be submitted in the formats previously described, then coded within the computer for internal use. For a question, all autopsy reports satisfying the question criteria would be output into a file, as is presently done, and then this coded file would be decoded and converted back to English text for output to the user. The net result would be approximately 40% less storage but more processing.

BIBLIOGRAPHY

1. Bahn, R.C., Schmit, R.W., and Lutz, T.D. "Potential Uses of a Digital Computer in the Section of Experimental and Anatomic Pathology", Mayo Clinic Proceedings, Vol. 39, No. 11, November 1964, pp. 830-834.

2. Bahn,R.C., Schmit, R.W., and Young, G.G. "An Information Retrieval System for Research Associated with the Postmortem Examination", Mayo Clinic Proceedings, Vol. 39, No. 11, November 1964, pp. 835-840.

3. Becker, H. "Befunddokumentation in der Pathologie: Erfahrungen mit einer Mashinenlochkartei", Methods of Information in Medicine, Vol. 4, No. 1, March 1965, pp. 30-35.

4. Becker, J. and Hayes, R.M. Information Storage and Retrieval: Tools, Elements, Theories, New York: John Wiley & Sons, 1963.

5. Carpenter, H.M. "A System for Storage and Retrieval of Data from Autopsies", The American Journal of Clinical Pathology, Vol. 38, No. 5, November 1962, pp. 449-467.

6. Committee on Nomenclature and Classification of Disease, Systematized Nomenclature of Pathology, Chicago: College of American Pathologists, 1965.

7. Crocker, D.W. "Preparation of Anatomic Pathology Data for Automation", Pathology Annual, Vol. 4, pp. 31-42, New York: Appleton-Century-Crofts, 1969.

8. Jacob, W. "Basic-Dokumentation in der Pathologie", Methods of Information in Medicine, Vol. 6, No. 4, October 1967, pp. 166-173.

9. Jacobs, H. "A Natural Language Information Retrieval System", Methods of Information in Medicine, Vol. 7, No. 1, January 1968, pp. 8-16.

10. Lamson, B.G. "Storage and Retrieval of Uncoded Tissue Pathology Diagnosis in the Original English Free-Text Form", Proceedings of the 7th IBM Medical Symposium 1965, pp. 411-426.

11. Lamson, B.G. and Dimsdale, B. "A Natural Language Information Retrieval System", Proceedings of the IEEE, Vol. 54, No. 12, December 1966, pp. 1636-1640.

12. Lindberg, D.A.B. The Computer and Medical Care, Springfield: Charles C. Thomas Co., 1968.

13. Marron, B.A. and DeMaine, P.A.D. "Automatic Data Compression", Communications of the ACM, Vol. 10, No. 11, November 1967, pp. 711-715.

14. Paplanus, S.H., Shepard, R.H., and Zvargulis, J.E. A Computer-Based System for Autopsy Diagnosis Storage and Retrieval Without Numerical Coding", Laboratory Investigation, Vol. 20, No. 2, February 1969, pp. 139-146.

15. Payne, L.C. An Introduction to Medical Automation, Toronto: J.B. Lippincott Co., 1966.

16. Pratt, A.W. and Thomas, L.B. "An Information Processing System for Pathology Data", Pathology Annual, Vol. 1, pp. 1-21, New York: Appleton-Century-Crofts, 1966.

17. Reid, W.D. and Heaps, H.S. "Compression of Data for Library Automation", presented at Automation in Libraries 1971, sponsored by The Canadian Association of College and University Libraries, Committee on Library Automation, Vancouver, June 1971.

18. Smith, J.C. "Basis for Automated Encoding of Anatomical Diagnoses", presented at International Academy of Pathology 58th Annual Meeting, 1969, abstract in Laboratory Investigation, Vol. 20, No. 6, June 1969, pg. 603.

19. Smith, J.C. and Melton, J. "Automated Retrieval of Autopsy Diagnoses by Computer Technique", Methods of Information in Medicine, Vol. 2, No. 3, July 1963, pp. 85-90.

20. Smith, J.C. and Melton, J. "Manipulation of Autopsy Diagnoses by Computer Technique", Journal of the American Medical Association, Vol. 188, No. 11, June 15, 1964, pp. 958-962.

21. Smith, J.C. and Melton, J. Data Control for Anatomic Pathology, Cleveland: Institute of Pathology, Western Reserve University, 1965.

22. Smith, J.C. and Melton, J. "Mechanized Coding, Storage, and Recall of Autopsy Diagnoses", presented at International Academy of Pathology 54th Annual Meeting, 1965, abstract in Laboratory Investigation, Vol. 14, No. 5, May 1965, pg. 564.

23. Thiel, L.H. and Heaps, H.S. Computer Search of Chemical Titles at the University of Alberta, Edmonton: Department of Computing Science, The University of Alberta, 1968.

APPENDIX

Program Output

The following pages contain output listings of a two-question batch. The first question demonstrates the abbreviated output option; the second the full output option.

```
*****************************************
*                                       *
*    S E A R C H   Q U E S T I O N      *
*                                       *
*****************************************

QUE     MARVIN ERAUCE
AND R   C
AND AL 090
NCT FE 70-0048
 NOR AE 000

*** ANSWER REFLECTS DATA BASE AS OF': 1971/08/06
```

AEBREVIATED RESULTS

101

DCE,JOHN	70-0046
SMITH,MARY	70-0052

```
*****************************************
*                                       *
*      S E A R C H    Q U E S T I O N   *
*                                       *
*****************************************
```

QUE   MARVIN BRAUDE. THIS TESTS FULL CUTPUT FORMAT
AND R   C
AND AL 090

*** ANSWER REFLECTS DATA BASE AS OF : 1971/08/06

103

THE UNIVERSITY OF ALBERTA HOSPITAL AUTOPSY SERVICE

PAGE 1

MARVIN FRAUCE. THIS TESTS FULL OUTPUT FORMAT

DOE,JOHN ,70-0046 VERSION DATE: 1971/06/03

GENERAL INFORMATION

AGE: 033 SEX: M RACE: C UHID: 314146
BIRTH: ADMIT: 1970/02/03 DEATH: 1970/02/09 AUTOPSY: 1970/02/10
ELAPSE: 16H COMPLETENESS: COMPLETE
RESIDENT: DR. CHERNEKOFF STAFF: DR. HAIN SERVICE: DR. WILLIAMS
CLINICAL DIAGNOSIS: ACUTE HEPATIC FAILURE

PATHOLOGICAL DIAGNOSIS (BODY PART,CONDITION,CONDITION,CONDITION)

OPERATION	CHOLECYSTECTOMY	1970/01/08	R70-291
	COMMON-BILE-DUCT	1970/01/08	R70-291
	T-TUBE	1970/01/08	
SKIN	JAUNDICE	TRANSIENT	
WOUND	SECONDARY	HEMORRHAGE	
SUBDIAPHRAGMATIC	RIGHT	HEMORRHAGE	ORGANIZING
OPERATION	LAPAROTOMY	1970/01/16	
	GASTROSTOMY	1970/01/16	
SKIN	JAUNDICE	SEVERE	
LIVER	NECROSIS	ACUTE	DIFFUSE
	RECENT		
OPERATION	LAPAROTOMY	1970/02/06	
PERITONEUM	PERITONITIS	EXTENSIVE	EFFUSION
ABDOMEN	UPPER	ADHESIONS	FIBRINOUS
	MODERATE		
PANCREAS	PANCREATITIS	INTERSTITIAL	ACUTE
OMENTUM	NECROSIS	FAT	EXTENSIVE
MESENTARY	NECROSIS	FAT	EXTENSIVE
SPLEEN	SPLENITIS	ACUTE	
KIDNEYS	NEPHROSIS	CHOLEMIC	BILATERAL
BLOOD	UREMIA		
ESOPHAGUS	LOWER	MUCOSA	HEMORRHAGIC
LUNGS	CONGESTION	EDEMA	
RT LUNG	BRONCHOPNEUMONIA	HYPOSTATIC	
OPERATION	TRACHEOSTOMY	1970/02/09	
ANKLES	EDEMA		
AORTA	ABDOMINAL	ATHEROSCLEROSIS	MINIMAL
URETERS	BIFID	BILATERAL	

MEASURES (PART,MEASURE(CM,GM,ML))

BODY LENGTH	000171.000	RT PLEURAL CAVITY	000200.000
LT PLEURAL CAVITY	000800.000	PERICARDIAL CAVITY	000050.000
THYROID	000032.000	LT LUNG	000520.000
RT LUNG	000790.000	HEART	000385.000
LT VENTRICLE	000008.000	RT VENTRICLE	000005.000
MITRAL VALVE	000010.000	AORTIC VALVE	000008.000
TRICUSPID VALVE	000011.000	PULMONARY VALVE	000008.000
SPLEEN	000280.000	LT KIDNEY	000270.000
RT KIDNEY	000240.000	PANCREAS	000065.000
LT ADRENAL	000010.000	RT ADRENAL	000010.000
LIVER	001720.000	BRAIN	001510.000

THE UNIVERSITY OF ALBERTA HOSPITAL AUTOPSY SERVICE

MARVIN FRAUCE. THIS TESTS FULL OUTPUT FORMAT

DOE, JOHN 70-0046 VERSION DATE: 1971/06/03

DRUGS AND OTHER THERAPY (THERAPY,DOSAGE(ML,GM),RATE,ROUTE,START,END / REASON ADMINISTERED / EFFECT)

HALOTHANE				1970/01/08	1970/01/08
ANESTHETIC					
TRANSIENT	JAUNDICE	HEMORRHAGE			
HALOTHANE				1970/01/16	1970/01/16
ANESTHETIC					
SEVERE	JAUNDICE				
PREDISONE				1970/01/26	1970/01/31
JAUNDICE					
AMPICILLIN			IV	1970/02/02	
PERITONITIS					
SERUM POTASSIUM	2.5 MGM				
POTASSIUM CHLORIDE				1970/02/03	
HYPOKALEMIC					
M-O-FLURANE				1970/02/06	1970/02/06
ANESTHETIC					
LOCAL ANESTHESIA				1970/02/09	1970/02/09
ANESTHETIC					
DETERIORATED					

MEDICAL HISTORY (DATE,SEQUENCE,FINDING,FINDING,FINDING)

1969/10/00	- 1	OBSTRUCTIVE	JAUNDICE	
1970/01/08	- 1	CHOLECYSTECTOMY	COMMON-DUCT	EXPLORATION
	- 2	T-TUBE	INSERTION	
	- 3	STONES	GALLBLADDER	COMMON-BILE-DUCT
	- 4	POST-OPERATIVE	TRANSIENT	JAUNDICE
	- 5	T-TUBE	CHOLANGIOGRAM	NORMAL
1970/01/16	- 1	LAPAROTOMY	FOR	SECONDARY
	- 2	HEMORRHAGE	GASTROSTOMY	
1970/01/22	- 1	MORE	JAUNDICED	
1970/01/27	- 1	SERUM-BILIRUBIN	41 MGM/100ML	
	- 2	PROTHROMBIN	38 SEC	
	- 3	CONTROL	14 SEC	
1970/01/30	- 1	BUN	46 MGM	
	- 2	GLUCOSE	155 MGM	
	- 3	BILIRUBIN	41 MGM	
1970/02/02	- 1	PERITONITIS	APPARENT	
1970/02/03	- 1	TRANFERRED	EDSON	UAH
	- 2	BLOOD	PRESSURE	130/90
	- 3	PULSE	100/MIN	
	- 4	TEMPERATURE	36.3 C	

MEDICAL HISTCRY - CONTINUED

	- 5	DEEPLY	ICTERIC	
	- 6	ABCCMEN	DISTENDED	
	- 7	1+	PEDAL	EDEMA
1970/02/04	- 1	FCTASSIUM	2.1 MEQ/L	
	- 2	WHITE	CCUNT	1600
	- 3	SHIFT	LEFT	
	- 4	4+	BILE	URINE
	- 5	REMAINED	HYPCKALEMIC	
	- 6	SERUM-AMYLASE	12 UNITS	
1970/02/05	- 1	BUN	88 MGM/100ML	
	- 2	TCTAL	BILIRUBIN	25.1 MGM/100ML
	- 3	CCNJUGATED	BILIRUBIN	10.7 MGM/100ML
1970/02/06	- 1	LAPAROTCMY	DRAINAGE	GREEN
	- 2	ABCCMINAL	FLUID	
	- 3	CULTURE	T-TUBE	GREW
	- 4	STAPHYLOCCCCUS-AUREU		
1970/02/07	- 1	BLCCD	SUGAR	344 MGM%
	- 2	URINE	2520 CC/24HR	
1970/02/08	- 1	URINE	2350 CC/24HR	
	- 2	URINE	AMYLASE	235 UNITS/24HR
1970/02/09	- 1	BUN	80 MGM%	
	- 2	SERUM-CREATININE	2.2 MGM%	
	- 3	BLCCD	GLUCOSE	543 MGM%
	- 4	TRACHEOSTOMY	1300 HR	
	- 5	CCNDITICN	DETERIORATED	
	- 6	DIED	1845 HR	

THE UNIVERSITY OF ALBERTA HOSPITAL AUTOPSY SERVICE

MARVIN BRAUDE. THIS TESTS FULL OUTPUT FORMAT

JONES,BABY GIRL 70-0048 VERSION DATE: 1971/03/23

GENERAL INFORMATION

AGE: 000 SEX: F RACE: C UFID:
BIRTH: 1970/01/31 ADMIT: 1970/01/31 DEATH: 1970/01/31 AUTOPSY: 1970/02/05
ELAPSE: 5D COMPLETENESS: COMPLETE
RESIDENT: DR. GEERE STAFF: DR. BAIN SERVICE: DR. RITCHIE
CLINICAL DIAGNOSIS: STILLBORN

PATHOLOGICAL DIAGNOSIS (BODY PART,CONDITION,CONDITION,CONDITION)

FETUS	MACERATED	STILLBORN	710 GM
FINGERS	ALL	ARACHNODACTYLY	
LUNGS	ATELECTASIS	PRIMARY	UNAERATED
ALVEOLI	PULMONARY	SQUAMES	ABUNDANT
TISSUES	DEGENERATION	ADVANCED	
BRAIN	LIQUIFACTION		
SKULL	COLLAPSED		

MEASURES (PART,MEASURE(CM,GM,ML))

BODY LENGTH	000037.000	BODY WEIGHT	000710.000
PLACENTA	000310.000	UMBILICAL CORD	000005.000
LUNGS	000019.000	KIDNEYS	000010.000
LIVER	000028.000	PHARYNX	000008.000
SPLEEN	000010.000	HEART	000060.000

THE UNIVERSITY OF ALBERTA HOSPITAL AUTOPSY SERVICE

MARVIN BRAUDE. THIS TESTS FULL OUTPUT FORMAT

SMITH,MARY 70-0052 VERSION DATE: 1971/06/03

G E N E R A L   I N F O R M A T I O N

AGE: 083 SEX: F RACE: C UHID: 179639
BIRTH: ADMIT: 1970/01/30 DEATH: 1970/02/10 AUTOPSY: 1970/02/12
ELAPSE: 2D COMPLETENESS: COMPLETE
RESIDENT: DR. LICHTENFELD STAFF: DR. JAIN SERVICE: DR. WILLIAMS
CLINICAL DIAGNOSIS: METASTATIC CANCER LEFT RADICAL
MASECTOMY 1961

P A T H O L O G I C A L   D I A G N O S I S   (BODY PART,CONDITION,CONDITION,CONDITION)

LT BREAST	CARCINOMA	SIMPLEX	
RT PLEURA	CARCINOMA	ANAPLASTIC	MASSIVE
RT LUNG	CARCINOMA	ANAPLASTIC	
EPICARDIUM	CARCINOMA	ANAPLASTIC	
LIVER	CARCINOMA	ANAPLASTIC	
NODES	LYMPH	MEDIASTINAL	PERIAORTIC
	CARCINOMA	ANAPLASTIC	
SPLEEN	CARCINOMA	ANAPLASTIC	
SKULL	CARCINOMA	ANAPLASTIC	
VERTEBRAE	CARCINOMA	ANAPLASTIC	
RT FISTULA	PLEUROCUTANEOUS	INFECTED	STAPHYLOCOCCUS-AUREU
RT FEMORAL	THROMBOPHLEBITIS		
ILIAC	EXTERNAL		
VEINS	EPIGASTRIC	DEEP	INFERIOR
LT LUNG	THROMBOEMBOLUS	MICROSCOPIC	
LUNGS	CONGESTION	EDEMA	MILD
	PATCHY		
RT LUNG	PNEUMONITIS	CHOLESTEROL	SECONDARY
LIVER	MARKED	METAMORPHOSIS	FATTY
AORTA	ABDOMINAL	ATHEROSCLEROSIS	MODERATE
ARTERIES	CORONARY	ATHEROSCLEROSIS	MARKED
ATRIUM	SEPTAL	DEFECT	SMALL
UTERUS	LEIOMYOMA	HYALINIZED	CALCIFIED
PANCREAS	CYST	BENIGN	RETENTION
APPENDIX	ADJACENT	CYST	PERITONEAL
	SMALL		
OVARIES	CYSTS	SEROUS	BENIGN
HEART	ATROPHY	BROWN	LIPOFUSCIN
	PIGMENTATION		

M E A S U R E S   (PART,MEASURE(CM,GM,ML))

BODY LENGTH	000150.000	BODY WEIGHT	000040.000
EYES	000000.300	LT LUNG	000380.000
RT LUNG	000150.000	HEART	000240.000
RT VENTRICLE	000000.300	LT VENTRICLE	000001.300
MITRAL VALVE	000000.300	AORTIC VALVE	000007.000
TRICUSPID VALVE	000011.000	PULMONARY VALVE	000007.500
SPLEEN	000080.000	LT KIDNEY	000085.000
RT KIDNEY	000065.000	UTERUS	000040.000
PANCREAS	000070.000	ADRENALS	000008.000

MEASURES - CONTINUED

LIVEF	00102C.0CC	BRAIN	001016.000

DRUGS AND OTHER THERAPY (THERAPY,COSAGE(ML,GM),RATE,ROUTE,START,END / REASON ADMINISTERED / EFFECT)

THIOTEPA		OR	1968/00/03	1968/00/04
PLEURAL	EFFUSICN			
LEUKOPENIA				
PILOCARBCN CROPS		OR	1953/00/03	1970/02/10
GLAUCCMA				

MEDICAL HISTCRY (CATE,SEQUENCE,FINDING,FINDING,FINDING)

1957/C0/00	- 1	FCST-INFLUENZAL	BELLS	PALSY
1961/10/27	- 1	RADICAL	MASECTOMY	CARCINOMA
	- 2	SIMPLEX	LT BREAST	
	- 3	AUXILLARY	LYMPH	NODES
	- 4	NEGATIVE	METASTATIC	DISEASE
1964/01/03	- 1	UNKNCWN	SPECIFIC	COLITIS
1968/00/03	- 1	PROGRESSIVE	DYSPNEA	
	- 2	13 LBS	WEIGHT	LCSS
1969/02/27	- 1	PLEURAL	EFFUSION	DISCOVERED
	- 2	1500 CC	GREENISH-YELLOW	DRAINED
1969/03/04	- 1	TUBF	INSERTED	2200 CC
	- 2	DRAINEC		
1969/03/22	- 1	TUBE	REMOVED	
1969/03/11	- 1	THIOTEPA	INJECTED	PLEURAL
	- 2	SPACE	RT LUNG	
	- 3	RELEASED		
1969/06/25	- 1	PLEURAL	EFFUSICN	1100 CC
	- 2	DRAINED		
1969/06/26	- 1	50 MGM	THIOTEPA	RIGHT
	- 2	PLEURAL	CAVITY	
	- 3	WHITE	CCUNT	1500-1600
	- 4	DRCP	1 MCNTH	
1969/07/28	- 1	STAPHYLOCCCCUS-AUREU	CULTURED	DRAINING
	- 2	SINUS		
	- 3	INTENSIVE	ANTIBIOTIC	TREATMENT
	- 4	DISCHARGEC		
1969/10/31	- 1	READMITTED	BUN	27
	- 2	DETERIORATEC	MENTAL	STATE
1970/01/30	- 1	READMITTED	CONFUSED	WEAK
	- 2	FELL	SUFFERING	SUPERFICIAL
	- 3	BRUISES		
1970/02/08	- 1	RT FCCT	EDEMATOUS	WARM
	- 2	INCONTINENT	URINE	FECES
	- 3	BLCCD	PRESSURE	DETERIORATING
1970/02/10	- 1	PRESSURE	60/40	MORNING

MEDICAL HISTORY - CONTINUED

- 2	CHEYNE-STOKES	RESPIRATION	AFTERNOON
- 3	DIED	2200 HR	